WO0178443

Publication Title:

EARSET COMMUNICATION SYSTEM

Abstract:

Abstract of WO0178443

A system and method for providing wireless commun 1161 ication which is controlled by voice recognition software running on a controller. The system includes an earset communicator and a Base Station that allows wireless communication between these elements. The earset communicator rests comfortably on the user's ear and is held in place by an earhook. The transceiver Base Station communicates with the earset communicator and connects to a host controller, such as personal computer ("PC") or a household product, and to a network interface such as an internet connection or phone line. Voice commands are used for many functions for controlling the system. The Base Station routes the earset microphone audio to the controller software for speech recognition and command processing. Speech recognition software on the controller interprets the voice command and acts accordingly. Data supplied from the esp@cenet database - Worldwide

---

Courtesy of http://v3.espacenet.com

This Patent PDF Generated by Patent Fetcher(TM), a service of Stroke of Color, Inc.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 October 2001 (18.10.2001)

PCT

(10) International Publication Number
WO 01/78443 A2

(51) International Patent Classification[7]: H04R 1/10

(21) International Application Number: PCT/US01/11069

(22) International Filing Date: 5 April 2001 (05.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/544,727 6 April 2000 (06.04.2000) US

(71) Applicant: **ARIALPHONE, LLC.** [US/US]; 3 Hawthorne Parkway, Vernon Hills, IL 60061 (US).

(72) Inventors: **PIRELLI, Thomas, R.**; 909 Oakhurst Lane, Riverwoods, IL 60015 (US). **PATEL, Sanjeev, D.**; 801 Salceda Drive, Mundelein, IL 60060 (US). **CYGNUS, Marc, W.**; 1111 Keswick Drive, Mundelein, IL 60060 (US). **SCHMID, Joseph, J.**; 3770 floara Vista Avenue, #1002, Santa Clara, CA 95051 (US). **WAGENER, Michael, E.**; 2572 N. Hickory Lane, Arlington Heights, IL 60007 (US).

(74) Agent: **SAMPSON, Matthew, J.**; McDonnell Boehnen Hulbert & Berghoff, Suite 3200, 300 S. Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/78443 A2

(54) Title: EARSET COMMUNICATION SYSTEM

(57) Abstract: A system and method for providing wireless communication which is controlled by voice recognition software running on a controller. The system includes an earset communicator and a Base Station that allows wireless communication between these elements. The earset communicator rests comfortably on the user's ear and is held in place by an earhook. The transceiver Base Station communicates with the earset communicator and connects to a host controller, such as personal computer ("PC") or a household product, and to a network interface such as an internet connection or phone line. Voice commands are used for many functions for controlling the system. The Base Station routes the earset microphone audio to the controller software for speech recognition and command processing. Speech recognition software on the controller interprets the voice command and acts accordingly.

TITLE: EARSET COMMUNICATION SYSTEM

FIELD OF THE INVENTION

The present invention relates to an earset communication system. The earset communication system includes a hands-free earset for use in Voice over Network (VoN) communication, voice dictation, control of a computer, and/or voice control of a number of additional functions (e.g., home entertainment and home automation).

BACKGROUND

Office communication products and systems have evolved significantly since the introduction of the telephone over 100 years ago. Today, one's home or office desk is frequently equipped with terminal devices such as computers, personal organizers, pagers and telephones allowing a user the ability to communicate by sending email, facsimiles, letters, telephone voice calls and voice messages. The development of these communication technologies has focused on providing the user with a choice of mediums for communication between terminal devices. However, along with the advantages resulting from the development of these communication mediums, the disadvantages of interfacing with these mediums has increased significantly.

A user today often must choose from a number of alternative communication mediums through specialized terminal devices such as telephone for voice, facsimile machine for a facsimile transmission of text and images, and computers for text, email, images, video, video conferencing and voice. Combinations of mediums and devices are also available, such as voice over IP, and data over phone lines providing the user an even greater variety of communication options. For example, using Voice over Internet Protocol (VoIP), Internet telephony may be combined with other modes of communication, such as video conferencing, and data or application sharing, giving a user tremendous power to communicate with others,

worldwide, at a fraction of the cost of conventional telephone systems. A disadvantage of typical VoIP systems, however, is that the user is tied to his or her computer when using the VoIP functionality

In contrast to traditional circuit switched telephone networks that are limited to transmitting voice or data within the conventional voice bandwidth, telephone switching systems are rapidly transitioning to packet-switched networks. In the packet-switched environment, information is transmitted over the network in short bursts of data, known as "packets." Packet-switched networks are generally more cost efficient than circuit switched networks because they require no call set-up time (resulting in faster delivery of traffic) and because users can efficiently share the same channel (resulting in lower cost).

The transmission of voice over a communication network may be referred to herein as Voice over Network ("VoN"). Voice over Internet Protocol ("VoIP") is used herein to refer to a specific form of VoN transmission: voice communication over packet-switched networks using the Internet Protocol. Today, VoIP is currently the most common implementation for VoN in the consumer market and is yet another selection available to a user for communicating more efficiently than ever before. However, increasing the number of communication mediums also increases the complexity of communicating because the user must decide which medium will be used and then interface with the appropriate device.

Although an office user has a wide variety of communication mediums to select from, the different systems typically each require their own user interface, resulting in increased complexity and ergonomic problems. In addition, an office user's work space must also provide a substantial amount of space for myriad devices, including for example a telephone-speakerphone, a computer keyboard, a mouse, a monitor, speakers and a microphone, a camera for voice and video over IP applications, and perhaps a personal digital assistant

("PDA"). Electrical connections are required for each product also creating messy cable nests. Additional office communication products may also be required or desired such as cellular telephones, pagers, printers, scanners, dictation machines and personal organizers, further increasing ergonomic problems by increasing options presented to the user, and further reducing valuable desk space as well.

The net effect of having multiple user-machine interfaces may actually result in reduced efficiency and productivity for the office user. As the number of communication devices increases, significant overlap in functionality and in hardware occurs. For example, a conventional telephone-speakerphone is largely redundant hardware for users with cordless telephones, cellular telephones and/or computer-based telephony devices. The telephone keypad and display is redundant with the computer keyboard and monitor since these functions may be combined to simplify the user-machine interface and reduce the required desk-top space. Existing devices also require the user to operate and maintain multiple terminal devices further contributing to ergonomic inefficiency. A typical modern office has poor ergonomics due to the incompatibility between these multiple devices and multiple interfaces, thereby requiring a user to learn to use and maintain each of them effectively.

Office communication device manufacturers have attempted to improve the user-machine interface by developing hands-free products and wireless communication systems in order to eliminate handsets and to promote freedom of motion. Although hands-free devices freed the user from having to hold a handset, these devices were limited to merely being extensions of a telephone handset. A user still has to manually control the communication device whether it is a telephone, answering machine, fax machine or a computer for sending email. Conventional cordless telephones utilize an RF link to provide wireless communication between the handset and the base station. However, conventional cordless

telephones are limited to establishing a wireless link between the handset and the base station with manual control interfaces.

Voice recognition systems were developed in order to convert speech into text based on the recognition of spoken words. For example, through the use of speech recognition software, a user does not have to use the computer keyboard in order to type text. Speech may be processed through a recognition algorithm resulting in the recognition of the word and the representation of the word as text or a computer display. These systems however have been largely limited to word processing applications.

Remote speaker and microphone systems are known in which a transceiver located in a headset is capable of establishing a link with a portable telephone. Such systems however, have several limitations. As an example, U.S. Patent No. 5,590,417, issued to Rydbeck, describes a wireless headset. In U.S. Patent No. 5,590,417, the contents of which are incorporated herein by reference, the wireless headset is worn on the user's head and receives and transmits a voice conversation to a portable telephone. One significant disadvantage of such a system is that the system cannot control functions such as dialing or searching for a telephone number without affirmative manual interface with the user. Specifically, a user still has to manually enter the phone number and initiate the call, usually by pressing a "Send" button. This has the disadvantage of distracting the user who may be operating a vehicle, PC, or another communication device.

A further disadvantage of such systems is that a user must manually mute the microphone or remove the headset in order to switch from speaking on the telephone to communicating on another device or speaking to others in the office without being overheard. The lack of the ability to command and/or control the communication device without manual intervention, therefore, limits the speed and efficiency of a user.

The communication and control challenges discussed above with reference to the home user may apply with equal force to the corporate setting. In the corporate office environment, the desirability of hands-free functionality has been demonstrated by the proliferation of headset products. A disadvantage of the known products, however, is that they still require a manual interface for even basic communication tasks like telephony. For example, with known headsets, to place a call, the user typically has to manually take a telephone off-hook and dial a number or similarly input commands through a computer keyboard or mouse. It would be desirable to free the user from the requirements of such manual interfaces.

Likewise, home automation applications are known for use with a personal computer. Again, however, the user is limited in the sense that the user must still use traditional computer interfaces, such as a keyboard or mouse to input commands. It would be desirable to provide the user with the freedom to control home automation functions without any type of manual interface with the computer.

No known system has successfully integrated the desired features into a system that provides a simple, intuitive control mechanism of the user's communication devices and mediums while also eliminating the need for redundant hardware and the requirements of numerous manual interfaces. It would therefore be desirable to have an improved communication system.

SUMMARY OF THE INVENTION

An object of the invention is to improve the efficiency and productivity of a home or office user. Productivity is improved by replacing existing conventional user-machine interfaces with a single convenient user-machine interface that is transparent to the user and responsive predominately to voice commands. By eliminating multiple inefficient conventional user-machine interfaces, the invention improves productivity by consolidating the functions of numerous communications interfaces to a single, hands-free, wireless communications interface. This also provides the advantage of eliminating training of the user to operate these various devices.

In accordance with a first aspect of the present invention, the user is provided with a hands-free, wireless earset that operates as a communication interface. The earset may provide control and/or communication functionality in accordance with voice commands issued by the user. For example, by way of the earset, VoN communication is enabled without requiring a manual control interface. For one embodiment, VoN communication includes VoIP.

Another aspect of the invention is to provide an earset communication system that provides control functionality using speech recognition software running on a microprocessor-based appliance. In one embodiment, the earset is coupled by air interface to a base station, which is capable of connecting to the microprocessor-based appliance (e.g., a PC, handheld computer, PDA, set top box, cable modem, and the like). In another embodiment, the base station connects directly to a PC (personal computer) and uses software running on the PC. Advantageously, the earset has the potential to control network functions, such as Internet connectivity, home entertainment functions (such as home TV, DVD, audio and/or video

systems and the like), home automation functions and the like, when connected to the microprocessor-based appliance.

The system includes an earset communicator and a base station that preferably allows wireless communication between these elements. The earset communicator allows hands-free and wireless operation of the communication system, thereby completely freeing the user from being confined to the desktop. The base station operates with a voice recognition and control programs in a controller, giving the user simple, fast and complete control of every communication capability through the controller, including for example control of telephony and data. Therefore, one embodiment of the invention combines the communication power and flexibility of a controller-communication system with control functionality via simple voice commands.

In accordance with another preferred embodiment, the earset communication system may be used for communication via Internet telephony or VoIP, voice browsing of the Internet, voice dialing and control management, voice dictation, PSTN telephony, and/or home control functions. The system allows the user to access files, review paperwork, work on the computer, and handle other office or home related activities without being tied to the desk because the earset has no tethering wires.

In accordance with another aspect of the invention, the earset is a lightweight battery powered device having a noise canceling microphone or microphone array. This earset has an advantage over existing headsets with long boom microphones because the microphone is located outside the user's field of vision so that the user can work without distraction, converse face-to-face, and even drink a beverage while using the phone or PC without having to move or remove the earset. The earset preferably allows the user to converse on a call, command other electronic devices, use both hands to type or perform other functions, get up

and move around the entire home or small office, all without the need to detach wires, remove the earset or carry around a cordless telephone.

In accordance with another aspect of the invention, the system automatically dials stored telephone numbers based on a voice command. Thus, no phone numbers need to be remembered by the user, and no digits are required to be manually dialed.

In accordance with yet another embodiment, the earset communicator interface provides tremendous advantages while in an automobile or involved in other activity that requires the use of both hands. A user may operate a mobile telephone, computer or other peripheral device in a hands-free mode. In one embodiment of the earset communication system, the earset device communicates with a PDA, which is in turn connected to a network that is capable of supporting voice communication. Another embodiment of the earset communicator system utilizes the base station to communicate with both the earset communicator and the wireless telephone network.

In accordance with yet another preferred embodiment, voice commands are used for all control functions. Voice recognition software allows the user to interact with, for example, a computer via spoken commands to initiate a VoIP call. The system uses voice recognition for control functions such as placing phone calls and answering phone calls. Voice recognition software may also be utilized in conjunction with commands received through the earset to perform other functions, such as checking schedules and appointments, controlling functions for audio, video, lighting, HVAC,(Heating Ventilation Air Conditioning), motorized windows and doors, etc., voice browsing of the Internet, voice dictation, and integration with existing 3rd party software to create unique vertical applications.

The earset preferably rests comfortably on the user's ear and is held in place by an earhook. A transceiver base station communicates with the earset via a wireless link. In

accordance with a preferred embodiment, the earset communicator is extremely lightweight (approximately 28 grams, or 1 ounce) so that it may comfortably be supported entirely by the user's ear, without the need for an over-the-head band.

BRIEF DESCRIPTION OF THE DRAWINGS

The preferred embodiments of the present invention are illustrated by way of example, and not limitation, in the figures of the accompanying drawings in which:

Figure 2 illustrates the earset;

Figure 3 illustrates the inner side of the earset;

Figures 4A & 4B illustrate two rear-view embodiments of the earset;

Figure 5A is a functional block diagram of one embodiment of the earset communication system that illustrates audio flow information and Figure 5B further illustrates a plurality of types of the network interface shown in Figure 5A;

Figure 6 is a functional block diagram illustrating a VoN audio interface in the microprocessor-based appliance shown in Figures 5A and 5B;

Figure 7A illustrates a block diagram of a preferred embodiment of the earset communication system including hardware and software system components;

Figure 7B illustrates an alternative embodiment of the earset communication system in which the network interface is incorporated into the base station;

Figure 8A illustrates the analog-to-digital and digital-to-analog conversions in the base station portion of Figure 7A and Figure 8B illustrates an alternative embodiment of that portion of the base station eliminating the analog portion of the path;

Figure 9 illustrates generalized software flow for handling the issuance of a command by the user;

Figure 10 illustrates the software flow for making a call;

Figure 11A illustrates the software flow for making a call where the user provides the called party's name and location;

Figure 11B illustrates the software flow for making a call where the user provides only the called party's name;

Figure 11C illustrates the software flow for requesting the voice agent;

Figure 12 illustrates the software flow for retrieving schedule information

Figure 13 illustrates the software flow for control of home entertainment functions.

Figure 14 illustrates the generalized software flow diagram for home automation functions.

Figure 15 illustrates a system for utilizing the earset and base station in a Voice over Network implementation.

DETAILED DESCRIPTION OF THE PRESENTLY PREFERRED EMBODIMENTS

An earset communication system and method of using such a system are described with reference to the figures introduced above. As shown in Figure 1, the earset communication system includes three main components: a wearable transceiver, hereinafter referred to as the earset or earset communicator 10; a transceiver base station 20 having an interface to a microprocessor-based appliance 30; and a microprocessor-based appliance 30. The microprocessor-based appliance 30 preferably includes at least one network interface, such as a network card or modem for access to the Internet or a corporate network and a PSTN telephony interface, for example a voice modem or similar device for access to the PSTN. Such similar devices for access to the PSTN include, for example, the PhoneRider by MediaPhonics or the Internet Line Jack by QuickNet. As described further below, the network interface may alternatively be incorporated into the base station 20.

The microprocessor-based appliance 30 preferably utilizes voice recognition software and communication software modules to interface with a communication medium. The

microprocessor-based appliance 30 may be, for example, a personal computer, a server, a PDA, a set top box, a cable modem, a handheld computer, or a web browsing kiosk. Media devices and other household controllers often are processor controlled, and therefore are capable of being integrated into the earset communication system. The microprocessor-based appliance 30 may utilize any type of computer architecture including conventional microprocessors and neural networked processors.

As described further below, the network interface 300 provided by the microprocessor-based appliance 30 couples the base station 20 to a network capable of supporting voice (e.g., the Internet, corporate intranets, corporate networks, the PSTN and the like). In accordance with a preferred embodiment, the network is a packet-switched network that supports VoIP, Voice over ATM, Voice over Frame Relay, Voice over cable, Voice over DSL, and the like. The network may be a wired network, a wireless network, or a combination of the foregoing. The network may be a local area network (LAN), but for communication applications will more typically be a wide area network (WAN), a combination of WANs, the Internet or the PSTN.

The microprocessor-based appliance 30 includes a read-only memory (ROM) structure, a random access memory (RAM) structure, associated data and address buses, and a port for coupling the microprocessor-based appliance 30 to the base station 20. In accordance with a preferred embodiment, the port that couples the microprocessor-based appliance 30 to the base station 20 is a Universal Serial Bus ("USB") port. Other types of wired or wireless connection may alternatively be used. In addition, the microprocessor-based appliance 30 is preferably a personal computer. Those skilled in the art of communications will recognize, however, that the microprocessor-based appliance 30 may alternatively be a handheld computer, a PDA, a set top box, a server, a cable modem, a web browsing kiosk or the like.

As used herein, the phrase web browsing kiosk refers to an appliance, which includes the microprocessor-based appliance 30 structure recited above, or equivalents thereto, that is specifically adapted for browsing the Internet.

The Earset

The earset 10 preferably includes an audio transducer 53, a speaker 52 and a microphone 50, as shown in Figures 7A and 7B. The audio transducer 53 may be used for ringing or other similar paging or notice type functions. Preferably the audio transducer 53 is capable of generating a tone that is loud enough to notify the user of an incoming call, page or the like. Alternatively, the speaker 52 may provide the notice-type functions of the audio transducer 53, although this is less preferable because volume limitations on the speaker 52 may prevent the user from hearing the ringing or paging tone when the earset 10 is not present on the users' ear. In accordance with an alternative embodiment, the user may hear audio from a speakerphone (not shown) instead of an audio transducer 53 or speaker 52.

Figures 2 and 3 illustrate an exemplary form of the earset 10. The earset 10 is designed to be worn comfortably on the user's ear. As illustrated in Figure 3, the speaker 52 extends from the earset 10 and is configured to be inserted into the user's ear. The speaker may be surrounded by gel and/or foam to improve comfort and fit of the earset 10. Alternatively, the earset 10 may be carried by the user.

Unlike a headset, the earset 10 is preferably extremely lightweight (approximately 30 grams, or 1 ounce) so that it may comfortably be supported entirely by the user's ear. The earset 10 is supported upon the user's ear by an earhook 14, as shown in Figures 2 and 3. The earhook 14 not only stabilizes the earset 10 on the user's head when worn on the ear but also orients the microphone 50 for reception of commands spoken by the user. This earhook 14 may be connected to the earset device 10 via a thermal plastic ring which has notched detents

for repeatable positioning. The earhook 14 may be made out of a plastic or flexible wire so it can mold to fit each ear comfortably. When the earset 10 is not worn on the ear, a lightweight earhook/speaker is plugged into a 2.5 mm jack which is located between the optional battery charging and parking contacts which are shown in Figures 4A and 4B.

As shown in Figures 2 and 3, the microphone 50 is mounted in a cavity at an end of the earset 10 that is distal from the earhook 14. For the embodiment shown in Figures 2 and 3, the microphone 50 is housed in an adjustable mini-boom. The microphone 50 housing is preferably acoustically insulated to minimize coupling of unwanted mechanical noise. The microphone signal line is preferably electrically shielded to prevent the coupling of unwanted RF energy. The use of the mini-boom, or equivalently—the extension of the length of the earset toward the lip plane, is required for the high signal-to-noise ratio demanded by currently available voice recognition software. From the standpoint of the user, and for simplification of the mechanical design of the earset 10, it would be preferable to eliminate the mini-boom and to instead simply mount the microphone 50 directly to the earset at a greater distance from the lip plane. It is envisioned that, as speech recognition software improves and the noise background therefore becomes less pertinent, the mini-boom may be eliminated from the earset. Other noise cancellation techniques known to those skilled in the art, such as the use of a noise canceling microphone array, may be used as an alternative to the mini-boom, or in conjunction with the mini-boom to enhance audio quality.

The microphone 50 is preferably a miniature, passive noise canceling electret element with a cardioid response pattern. The mini-boom is pivotally attached to the body of the earset 10 to allow the mini-boom to pivot away from the major axis of the earset 10. Preferably, the mini-boom may pivot up to approximately 20° away from the major axis. When the earset 10 is worn by the user, the end of the mini-boom locates microphone to the

side of the user's mouth and approximately even with the lip plane while keeping the microphone out of the puff stream.

In alternative embodiments, the mini-boom is eliminated and the single microphone 50 and mini-boom are replaced by a microphone array with an associated DSP system that is programmed to reduce background noise and echoes. It is also envisioned that speech recognition software will in the future progress to the point where the noise cancellation techniques described above are not required. The obverse of the microphone 50 may be ported to enhance passive noise cancellation. Either active or passive noise cancellation techniques may be used. For example, an array of microphones may be used with a adaptive combiner to select a weighted group of microphone signals to provide the lowest noise and therefore the highest signal to noise ratio.

The speaker bud 114 shown in Figure 3 preferably extends from the body of the earset 10 and is covered by an acoustically permeable foam cap, which acts as a cushion to prevent the convex covering of the speaker bud 114 from irritating the ear. The speaker 52 is optimally capable of reproducing sound in the voice audio frequency band. The convex shape allows it to self-seat, centering upon the ear canal (in the Concha), with minimal to no adjustment, when placing the earset 10 upon the ear.

The earset 10 may be powered by a lightweight rechargeable battery 54, such as a Lithium-Ion Polymer battery. Other types of rechargeable batteries may alternatively be used. Without limiting the invention, a battery having the following characteristics is acceptable for the present application, although other batteries may alternatively be used. The weight of the battery may be approximately 7 grams or less. The dimensions may be approximately a width of 20 mm by length of 50 mm by a depth of 5 mm. The battery may have an approximate capacity of 250 mAH or more, and be capable of powering the earset for more than 2 hours.

The approximate battery voltage may be from 3.3 V to 4.1 V with an approximate nominal voltage of 3.8 V. Future improvements in battery performance including increased volumetric energy density and increased gravametric energy density may also be utilized. The battery 54 may be encased in a plastic pack that is mounted on the side of the earset 10 from the back as shown in Figures 4A and 4B. Preferably, the earset 10 includes battery charging/power contacts that are connected to the battery pack internally, i.e. through the earset, and the base station 20 includes mating contacts for charging the battery when the earset 10 is not is use. Alternatively, the battery may be removed from the earset 10 for charging, such as in a charging stand that may be incorporated into the base station 20.

The battery 54 is preferably located as close to the ear as possible to keep the center of gravity of the earset 10 nearest the center of the ear, and to be positioned to balance the earset 10. For power management purposes, the earset communicator 10 may normally be in a "sleep," or inactive, status in which most of its systems and components are powered down.

In accordance with one preferred embodiment, the earset 10 also includes a set of parking contacts as illustrated in the alternative embodiment of Figures 4A and 4B. When the earset 10 is in contact with the base station 20, and the parking contacts engage mating contacts on the base station 20, an identification code, which is commonly associated with a radio transceiver chipset within the earset 10, is sent by the earset 10 to the base station 20. In this manner, the base station 20 becomes associated with a particular earset 10. In other words, the base station 20 will communicate with the proper earset 10 even in an environment in which numerous earsets 10 are transmitting command signals.

For those situations where the user does not wish to wear the earset 10 on their ear, the earset may be provided with a separate speaker/microphone which can be plugged into an optional 2.5mm jack at the rear of the earset, as shown in Figure 4A. When the earset 10 is

inserted into the jack, the audio is diverted from the internal microphone 50 and speaker 52 to a connected wired speaker/microphone or speakerphone. Using a special clip that may attach for example to the speaker bud, the user may then attach the earset to his or her shirt, or wear it with a lanyard around their neck. Since the wired microphone/speaker typically weighs only 1/8 of an ounce (3.5 grams), this may be a more comfortable arrangement for some users.

Figure 7A illustrates a block diagram of a preferred embodiment of the earset communication system. Figure 7B illustrates an alternative embodiment. The system uses an RF link 180 to provide hands-free operation between a self-contained compact earset 10 and a base station 20, which has interfaces to a microprocessor-based appliance 30 and a communication network 300. The earset communicator 10 comprises a radio frequency transceiver system 62, 60 for wireless radio frequency between the earset 10 and the base station 20. The radio transceiver 60 is preferably a 900 MHz Digital Spread Spectrum Transceiver Model No. RF105, which is commercially available from Conexant Systems, Incorporated of Newport Beach, CA. This chipset, for example, will automatically select one of 40 available channels. By selecting the channel with the least interference and by utilizing DSS (Digital Spread Spectrum) technology, the system is interference tolerant. Radio transceiver 60 also preferably includes a Conexant 900 MHz Class AB RF Power Amplifier Model No. RF106 which provides a communicating range of approximately 250 feet (76 meters). The earset Codec 58 is preferably a Hummingbird 100-pin ASIC + CODEC (single chip) Model No. RSST7504 or equivalent. The base station Audio Processor 272 is preferably a 144 pin Hummingbird ASIC Model number RSST7504, and the base station CODEC 224 is preferably a 32 pin Hummingbird CODEC Model number 20415. The RF antenna 56 may reside within the plastic enclosure of the earset 10 provided the antenna 56

meets the minimum fractional wavelength requirements of the transmit frequency. The antenna 56 may be positioned along the outer edge of the plastic earset case.

Alternatively, transceivers 62, 60 may each be a 2.4 GHz spread spectrum transceiver system such as is available from Siemens Electronics, or a 900 MHz chipset such as offered by Rockwell/Conexant (as previously discussed), or an Ericsson bluetooth chipset, Model No. PBA 313 1/2, or any other chipset that supports wireless communication. Typically these chipsets are based on a full duplex analog, CDMA or TDMA technology formats. Chipsets from other manufacturers may alternatively be used, provided their air interface specifications provide high quality voice and security. One skilled in the art is capable of identifying commercially available components for the air interface in the system and would also recognize other substitute chipsets. An advantage provided by the 900 MHz and 2.4 GHz chipsets, however, is that they provide the earset 10 with a substantially longer usable range than is available from known headset arrangements.

The output of the earset radio receiver 60 is connected through the ASIC 108 to an amplifier in the CODEC 58 where the output portion of the audio circuit will drive the speaker 52. The output level of the signal sent to the earset speaker 52 is controlled digitally by the Hummingbird chip 108.

A tone may be emitted from an internal audio transducer 53 to alert the user of a low battery state. In addition, an out of range tone may optionally be emitted by the internal audio transducer 53 when the earset 10 is not within the recognizable range of the base station 20. When the earset 10 cannot sense the base station 20, the earset 10 preferably emits a specific tone, for example, periodically every 10 seconds. The earset 10 will emit a repeating ringing tone, preferably via the audio transducer 53, to notify the user of an incoming call. When the voice agent needs to present the user with a call notification, the microprocessor-based

appliance 30 may send a signal to the base station 20, which in turn relays the signal to the earset 10 to begin the ringing tone. The user preferably may locate the earset 10 by activating a paging signal from the computer 30, or the base station 20 for the optional case in which the base station 20 includes a button for sending the paging signal. The earset 10 may emit a repeating paging tone cadence to allow the user to locate the earset 10.

The earset communicator 10 contains controls that allow the user to switch the earset 10 to an "on" or active state when use of the earset functions is desired or necessary, such as when answering an incoming telephone call. A single button, i.e. a command button 110, on the earset communicator 10 prompts the microprocessor-based appliance 30 that a voice command is imminent. As described further below, the user preferably receives, in response to the user depressing the command button 110, a configurable ready prompt through the earset internal audio transducer 53 from the microprocessor-based appliance 30. The ready prompt notifies the user that the system is preferably ready to receive a voice command. The ready prompt is stored on the microprocessor-based appliance 30 for example in a digital sound file format that allows the user to configure or record customized prompts. The earset internal audio transducer 53 may also be used to notify the user of system status such as incoming phone calls, low battery status, paging signals, and "out of range" warnings.

The Base Station

The base station 20 is the communications gateway between the microprocessor-based appliance 30 and the earset 10 in the earset communication system. Reference may be made to Figures 7A and 7B for block diagrams of the base station 20, wherein the preferred embodiment is illustrated in Figure 7A and an alternative embodiment is shown in Figure 7B. The base station 20 contains circuitry necessary to operate the earset 10. The base station 20 footprint is preferably small relative to a desktop. In accordance with a preferred

embodiment, the base station 20 is small enough to be conveniently used while traveling, such as with a laptop computer. An internal RF antenna 22 may be used in order to provide a more aesthetically pleasing appearance, however, an external antenna 22 may alternatively be used. Antenna diversity may be utilized to increase signal to noise ratio and decrease RF interference.

In accordance with a preferred embodiment, the transceiver base station 20 provides a USB interface 21 to the microprocessor-based appliance 30, having an associated memory structure. As previously noted, the microprocessor-based appliance 30 may be a personal computer ("PC"), PDA (personal data assistant), or other microprocessor-based device such as a set top box, cable modem, or other Internet device/appliance, or home control/automation system or other Internet services device. Other types of interfaces to the microprocessor-based appliance 30, such as RS-232, PCMCIA, Bluetooth or infrared, may alternatively be used.

Figure 8A illustrates a portion of the base station 20 hardware from Figure 7A and illustrates the form of the voice signal between the USB interface 21 and the Hummingbird ASIC 272. As shown in Figure 8A, the voice signal is digital, such as 16 bit, 8kHz linear PCM data, between the ASIC 272 and the CODEC 224, which then converts voice signals from the ASIC 272 into analog form. The voice signal is then digitized by the CODEC 282 and passed to the USB interface 21. The opposite conversions are made for signals traveling from the USB interface to the Hummingbird ASIC 272. The intermediate conversion to analog form allows the Hummingbird ASIC 272 and the USB interface 21 to operate using independent clocks. In an alternative embodiment in which the Hummingbird ASIC and the USB interface 21 operate on synchronized clocks, the intermediate conversion may be eliminated as shown in Figure 8B.

Preferably, the base station 20 draws power entirely from the USB connection 21 to the computer 30. Alternatively, the base station 20 may be powered from a DC power adapter connected to an AC power source, commonly known to those skilled in the art. This alternative power source may be required where the base station 20 provides battery charging capability as noted above. The base station 20 may be a standalone unit, or may attach directly to the microprocessor-based appliance 30. For example, where the microprocessor-based appliance 30 is a laptop computer, it may be desirable to mount the base station 20 to the laptop for ease of use during transit. For example, this permits the user to use the system for voice dictation while traveling.

As an alternative, the base station 20 may be incorporated into the microprocessor-based appliance 30, either by physically incorporating the base station 20 hardware into the appliance 30 form factor or, where the appliance 30 is already capable of supporting a wireless connection to the earset, by programming the appliance 30 to perform the base station 20 functions. For example, it is envisioned that personal computers, PDAs, cellular telephones and the like will include transceivers that support communication in accordance with the Bluetooth protocol. Those skilled in the art would be capable upon reviewing this document of adapting the earset 10 to interface with such appliances 30.

The base station 20 provides an interference-resistant, secure RF link for multiple earsets. In one embodiment, the system may support up to 8 earsets. If multiple earsets 10 are communicating simultaneously, they act as "Conference Call" units, working in the same manner as multiple wired telephones on a single line. The earset to base station range is preferably in excess of 75 meters in the presence of interference from structures such as walls and ceilings. The signal between earset 10 and base station 20 is preferably capable of

passing through a minimum of six standard wood stud and drywall walls, which are typical of residential construction.

The earset 10 has the ability to associate itself with a specific base station 20 when in the presence of multiple base stations within the reception area. For example, as described above, the earset 10 may include parking contacts that, as is known in the art of cordless telephones, allow the earset 10 and base station 20 to be logically mated. In the same manner, the base station 20 and earset 10 may be set up to use a particular encryption technology.

One skilled in the art can readily implement such a system based on the air interface standards used in the radio transceiver chipset for the air interface 180. For example, the manufacturers of 2.4 GHz or 900 MHz digital spread spectrum chipsets associate a p.n. (pseudo random) code for those chipsets based on CDMA technology and these chipsets are readily utilized in this system. This capability will allow multiple earsets or earset systems to function simultaneously. Because the earsets 10 may be logically mated with a base station 20, the system allows many earsets 10 be associated with a single base station 20, or alternatively allows numerous earset 10/base station 20 pairs to be operated within the same area.

Voice over Network Communication

A preferred embodiment of the present invention provides advantageous use of the earset 10 with Voice over Network (voice over IP, voice over ATM, voice over Frame Relay, voice over cable, voice over DSL, and the like) technology. In Figure 7A, the microprocessor-based appliance 30 includes a network interface 300 that is accessible to the earset communication system via the software shown. In the alternative embodiment of Figure 7B, the base station 20 includes a network interface 300, which may be a DAA or "Data Access Arrangement" where the interface is to the PSTN. The network interface 300

may be a connection which couples the appliance 30 (in the case of Figure 7A) or the base station 20 (in the case of Figure 7B) to a communication link such as a data service, Internet service, cable modem type service, or a conventional telephone network interface (also referred to as the "TelCo") 25. For example, the network interface 300 may connect directly to an Internet data service in order to provide VoN functionality in a consumer or home office environment. In a corporate application, the network interface 300 may connect to a LAN, WAN or corporate network.

Figure 15 illustrates a block diagram of an embodiment of the present invention for using the earset 10 in conjunction with VoIP software to make an Internet-based 310 VoIP call. The earset 10 may also be used within a corporate telecommunications enterprise 390 to make voice over network calls when integrated with a corporate VoIP (or any VoN) platform such as those offered by 3Com Corporation, Cisco Systems and others.

Following are three exemplary scenarios describing the use of the earset 10 in conjunction with VoIP software. In the first scenario, as illustrated in Figure 15, VoIP calls are made between the earset 10 and microprocessor-based appliance 30'.

1. Microprocessor-based appliance 30 is connected to the IP network (Internet) 310.
2. User speaks into the earset 10.
3. Voice is transmitted over the air interface 180 in a transmission to the base station 20.
4. Voice is transmitted (digital) via USB 21 to the microprocessor-based appliance 30.
5. Voice is transmitted to the IP client (software), as shown in Figure 6, on the microprocessor-based appliance 30.
6. Voice is converted into IP packets and transmitted through the network interface 36, shown in Figure 6, to the microprocessor-based appliance 30' via the Internet 310.

Note that microprocessor-based appliance 30 to microprocessor-based appliance 30' VoIP communications do not require a VoIP gateway service provider 320. There are a number of software packages (including the Internet Phone client software offered by ArialPhone LLC of Vernon Hills, Illinois, as well as Microsoft NetMeeting, Internet Phone by VocalTec, Cu-Cme and the like) that can be purchased or downloaded from the Internet that allow users to talk to each other using their microprocessor-based appliances 30 & 30' and VoIP.

In the second scenario, illustrated in Figure 15, VoIP calls are made between the earset 10 and telephone 380 or Corporate desktop equipment 390 via Centrex Service.

1. Microprocessor-based appliance 30 is connected to the IP network (Internet) 310.
2. User speaks into the earset 10.
3. Voice is transmitted over the air interface 180 in a transmission to the base station 20.
4. Voice is transmitted (digital) via USB 21 to the microprocessor-based appliance 30.
5. Voice is transmitted to the IP client (software), as shown in Figure 6, on the microprocessor-based appliance 30.
6. Voice is converted into IP packets and transmitted through the network interface 36, shown in Figure 6, to an IP Gateway 320 via the Internet 310.
7. The IP Gateway 320, in this scenario typically part of the telephone company central office, converts the IP voice packets to analog and forwards the packets to the Central Office switch 330.
8. Central Office switch 330 transmits analog voice to analog telephone 380 or to Corporate desktop equipment 390 via Centrex..

In the third scenario, illustrated in Figure 15, VoIP calls are made between the earset 10 and Corporate desktop equipment 390 via telephone company Central Office switch 330.

1. Microprocessor-based appliance 30 is connected to the IP network (Internet) 310.
2. User speaks into the earset 10.
3. Voice is transmitted over the air interface 180 in a transmission to the base station 20.
4. Voice is transmitted (digital) via USB 21 to the microprocessor-based appliance 30.
5. Voice is transmitted to the IP client (software), as shown in Figure 6, on the microprocessor-based appliance 30.
6. Voice is converted into IP packets and transmitted through the network interface 36, shown in Figure 6, to an IP Gateway 320 via the Internet 310.
7. The IP Gateway 320, in this scenario typically part of the telephone company central office, converts the IP voice packets to analog and forwards the packets to the Central Office switch 330.
8. Central Office 330 transmits to corporate PBX 370, or IP PBX 360 (in this case, there is an IP Gateway 350 between the CO (central office) 330 and the IP PBX 360 to convert analog voice into IP Packets).
9. PBX 370 or IP PBX 360 transmits voice to the corporate telecommunications network 390.

As an alternative to the use of a telephone company Central Office switch 330 in scenario three, the IP packets may be routed directly to an IP PBX 340 and delivered in IP form to the corporate desktop equipment 390. Those skilled in the art will recognize that the path in Figure 15 that will be used for communication with the corporate desktop equipment 390 in any particular case is dependent upon the corporate desktop equipment 390 hardware.

A method is described below with reference to Figure 6 for interaction between the earset 10 and the microprocessor-based appliance 30 to make VoIP calls. It should be

recognized that the same method applies to other VoN protocols simply by replacing the IP client with an appropriate client that supports the desired protocol.

When the earset 10 user is speaking:

1. The user speaks into the microphone 50 on the earset communicator 10.
2. The earset communicator 10 transmits the analog voice to the base station 20 over the air interface 180.
3. The base station transmits the analog voice to the microprocessor-based appliance 30 using a USB connection 21.
4. The USB audio driver 32 passes the voice to the IP Client application 34.
5. The IP client application 34 converts the analog USB voice to IP voice packets.
6. The client application 34 transmits the IP voice packets to the microprocessor-based appliance's 30 network interface 36, such as a card or modem.
7. The PC's network interface 36 transmits the IP voice packets over the Internet 310.

When the earset 10 user is listening:

1. The PC's network interface 36 receives IP voice packets and passes them along to the IP client application software.
2. The IP client application converts the IP voice packets to analog voice.
3. The USB audio driver 32 passes the analog voice to the base station 20 via a USB connection 21.
4. The base station 20 passes the analog voice to the earset communicator 10 over the air interface (i.e. using a wireless transmission) 180.

Corporate Voice over Network

The use of VoIP in the corporate environment results in a significant reduction in the cost associated with intra-office (branch to branch), and inter-office communications. The cost of intra-office communication can be broken down into: equipment, maintenance, and telephone charges. Equipment and maintenance costs are the primary areas of savings for inter-office communications. VoN technology can significantly reduce these costs in the following manner:

Equipment – Generally speaking, VoN equipment is less expensive than traditional telephone equipment. Additionally, with VoN technology voice traffic travels over the same network infrastructure as data traffic meaning there is no need to purchase and maintain a completely separate network to handle voice.

Maintenance – Because VoN technology utilizes the existing data network there is no need to maintain a completely separate voice network. Also, existing IS staff generally has the knowledge to support and maintain the existing data network so there is no need to hire and train duplicate staff to manage the voice communications component.

Telephone Charges – Because VoN communication technology uses the existing data network, there is no need to lease separate lines to handle voice traffic in the case that the branch offices each have connected telephone equipment. In the event each branch office is not connected, and is using service provided by a long distance carrier, the savings can be greater because all long distance charges for intra-office calls can be eliminated.

In a preferred embodiment, the earset communicator system and software may be integrated with the offerings of VoN providers to add significant functionality including: voice agent capability to create "Intelligent Dial Tone," voice dialing, voice access to all

telephony features (park, call, transfer, etc.) and voice mail, and integration with corporate contact management and collaboration systems (Microsoft Outlook, Lotus Notes, etc.).

For example, the earset communication system preferably includes a VoN telephony system to provide a highly convenient, highly functional alternative to the soft phone (computer software) or telephone handset hardware. The earset communicator 10 preferably supports functionality with both, VoN and traditional voice solutions. The embodiments disclosed do not preclude working with standard telephone services. All the telephone functions described in this section apply to any transport medum, however the physical transport medium in the case of VoIP is based on the Internet Protocol.

Consumer VoIP

Another embodiment of the earset communication system provides IP Telephony in the consumer market to provide free or greatly reduced cost of long distance and international telephone calls. One drawback consumers face when using VoIP to make telephone calls is the fact that they are tied to their computer in order to receive the lowest possible rates (PC to PC or PC to Phone calls). That is, they are forced to use soft phone functionality via a graphical user interface supplied by the VoIP provider. They also generally must use a speaker and microphone combination wired to the computer.

However, since the earset communication system includes a wireless connection to a microprocessor-based appliance 30, through the base station 20, users can make VoIP calls from anywhere in the home, allowing them to use the earset communication system in conjunction with a VoIP provider to make calls like they might otherwise make using a standard telephone handset. Another key advantage that the earset communication system adds to the VoIP platform is voice dialing, making the process of initiating and answering IP telephony calls extremely simple and convenient. Additional functionality accessible via the

earset communication system software, such as voice mail, call screening, and unified messaging, round out the VoIP offering and make the complete solution an improvement over the existing analog telephone.

The primary service providers in the consumer VoIP market are demonstrating that the potential from this technology is significant. Some of the current VoIP providers are: Net2Phone, (http://www.net2phone.com), PhoneFree (http://www.phonefree.com), and DialPad (http://www.dialpad.com).

To effectively use VoIP today, a consumer may utilize a high speed Internet connection like DSL or a cable modem (standard 33k – 56k dialup will also work, although the voice quality may be somewhat less than that of standard telephone service). One of the primary problems with using VoIP is the fact that the user is tied to their computer – a problem that earset communication system neatly resolves. In addition to VoIP functions, additional capabilities that are enhanced by the earset communication system include voice chat for instant messaging, and voice-based command and control applications.

Instant Messaging Users

With approximately 45 million users of AmericaOnline's AOL Instant Messenger (AIM), and approximately 50 million ICQ users, plus the users of Yahoo! Messenger, MSN Messenger, and others, the instant messaging market consists of a substantial user base. Some of these instant messaging products support voice conversation, while others only offer text-based chats. Today, all of these services require that users be at their computers to engage in a chat. Integrating the earset communication system into these products allows users to initiate, respond to, and engage in a voice-based chat via the instant messaging software from anywhere in the home. Even without such integration, the earset communication system enables the users of instant messaging software that supports voice conversations to do so in a

hands free manner while the user is moving freely throughout the home (although users will still have to initiate and answer the chat at the computer).

Telephone Communication

Telephone communication will now be described with reference to Figures 5A, 5B, 7A and 7B. Figure 5A is a functional block diagram of the earset communication system. As shown, the microprocessor-based appliance 30 includes an interface 21 for communicating with the earset 10 via the base station 20 and also includes a network interface 300 for coupling the earset 10 via the appliance 30 to a network 80 that supports voice communication. Figure 5B shows that the network interface 300 may include one or more of: a network connection, such as a connection to a LAN, WAN, the Internet and the like, and a connection to the PSTN, such as by a USB PSTN interface 46 or PSTN Telephony Interface 48. The software 31 shown in Figures 5A and 5B is further described in Figures 7A and 7B. The software modules shown in Figures 7A and 7B, other than the earset agent application 320, are well known to those skilled in the art and are widely available. The preferred earset agent application is commercially available as the Arial Voice Agent software, from ArialPhone LLC of Vernon Hills, Illinois.

Figure 7A further describes the preferred embodiment in which the microprocessor-based appliance 30 includes both a network interface or NIC and a PSTN Telephony Interface. In the alternative embodiment shown in Figure 7B, the microprocessor-based appliance's 30 PSTN Telephony Interface is replaced by the USB PSTN interface, which is illustrated as residing in the base station 20. As used herein, the PSTN Telephony Interface may be a voice modem, Dialogic D/41ESC, PhoneRider by MediaPhonics type boards, Internet Phone Jack by QuickNet type boards, or the like.

The network interface card 342 provides the interface for VoN communication, as described above with reference to Figure 6. Such cards are readily available from 3Com Corporation of Santa Clara, California, Intel Corp. of Santa Clara, California and others, and provide full-duplex capabilities. This interface is not utilized for PSTN telephony.

In accordance with the preferred embodiment of Figure 7A, the PSTN Telephony Interface in the microprocessor-based appliance 30 includes DTMF dialer circuitry that is capable of dialing a phone number transmitted from the microprocessor-based appliance 30 via its internal bus. The PSTN Telephony Interface may include Caller ID detection circuitry that is capable of passing a caller's telephone number and test string to the microprocessor-based appliance 30 via its internal bus. In addition, the PSTN Telephony Interface preferably provides to the microprocessor-based appliance 30 audio I/O support of 16-bit, 8-KHz PCM formats: unsigned linear, G.711. Preferably, a four conductor RJ-11 jack may be used to couple the PSTN Telephony Interface to a telephone line.

Preferably, the PSTN Telephony Interface also has full-duplex audio circuitry that is capable of taking a first audio stream from the telephone line and placing it on the internal bus of the microprocessor-based appliance 30. The earset agent application 320 in conjunction with the well known device and media streaming drivers is capable of taking the first audio stream from the internal bus and transmitting it to the earset 10 via the base station 20. In the same manner, the earset agent application 320 is capable of placing a second audio from the earset 10 via the base station 20 onto the internal bus. The PSTN Telephony Interface is capable of taking the second audio stream from the internal bus and placing it on the telephone line. For full-duplex communication, the first and second audio streams are processed simultaneously in the earset communication system.

As the user speaks telephony control commands into the earset 10, they are transmitted to the earset agent application 320 via the base station 20. In response, the earset agent application 320 issues appropriate telephony control commands, such as on-hook, digit dialing, off-hook, flash, conference, mute and the like, to the PSTN Telephony Interface via the internal bus of the microprocessor-based appliance 30. In addition, the full-duplex audio processing will allow the earset agent application 320 to record line or earset audio, and to communicate voice commands, play back PC audio to the line or earset 10. For example, the microprocessor-based appliance 30 is able to send earset control codes to the base station 20 to permit signaling and prompting to the earset 10 to perform a specific function.

In accordance with the alternate embodiment of Figure 7B, the base station 20 has DTMF dialer circuitry (not shown) that will be capable of dialing a phone number transmitted from the microprocessor-based appliance 30 via the USB connection 21. The base station 20 also may include Caller ID detection circuitry 23 that is capable of passing a caller's telephone number and test string via the USB connection to the computer 30. In addition, the base station 20 preferably provides to the microprocessor-based appliance 30 audio I/O support of 16-bit, 8-KHz PCM formats: unsigned linear, G.711. In terms of the telephone network interface, the base station 20 includes a USB PSTN interface 46. A four conductor RJ-11 jack may be used to couple the base station 20 via the USB PSTN interface 46 when connected to a telephone line.

In one embodiment, the base station 20 also has full-duplex audio circuitry that is capable of communicating the audio stream provided via the USB connection 21 to the microprocessor-based appliance 30. Using the USB connection 21, the microprocessor-based appliance 30 and base station 20 will communicate telephony control commands as well as full-duplex audio processing. This allows the earset agent application 320 via the

microprocessor-based appliance 30 to control functions such as on-hook, off-hook, flash, conference and mute. In addition, the full-duplex audio processing will allow the earset agent application 320 to record line or earset audio, and to communicate voice commands, play back PC audio to the line or earset 10. For example, the microprocessor-based appliance 30 is able to send earset control codes to the base station 20 to permit signaling and prompting to the earset 10 to perform a specific function.

During a conversation between the earset 10 and the network 80, as shown in Figure 5A, the microprocessor-based appliance 30 may send an audio message to the earset 10, for example to alert the user of a call waiting. The earset agent application 320 may communicate separately and simultaneously with both the local and remote parties when the parties are not communicating with each other. For example, the local party may perform an Internet look-up while the remote party receives a recorded music stream. In addition, where no one is available at the earset 10 (i.e. local user), the earset agent application 320 via the microprocessor-based appliance 30 may communicate with the remote party to prompt the remote party to leave a message.

Vertical Market Applications

The unique form factor of the earset communication system provides significant support for vertical market solution providers to offer new, highly differentiable services. Examples of vertical market services include:

Public Safety

Application to allow a public safety officer to interview incident witnesses and automatically fill out forms and reports via a voice based interface using the earset communicator system . The application may also allow the officer to make voice requests for information via a central computer or the Internet.

Utility Workers

Application to allow utility workers to make voice requests for specifications on equipment that they are currently working on. The information from the central computer or the Internet is requested and delivered via the earset communicator system .

Medical/Legal Service Providers

Application that allows voice dictation of case/procedure notes. The application may also allow the service professional to request and retrieve information via the earset communication system.

Software Interfaces

In accordance with a preferred embodiment, voice commands are used for all functions and control of the system. When a user activates the command button to issue a command, the base station 20 routes audio picked up from the earset microphone 50 to the microprocessor-based appliance 30, where speech recognition is applied to the input command signal and the command signal is processed. Speech recognition software on the microprocessor-based appliance 30 interprets the voice command as described in greater detail below with reference to the software flow figures. In accordance with a first embodiment, only commands are routed to the microprocessor-based appliance 30 and not audio during a conversation with another party. Once the user has issued the command to make a call, communication audio (i.e., the audio from a VoN conversation) is not picked up by the earset agent application 320. The reason for this is that it is not practical for a number of reasons to have the speech recognition software listen to an entire conversation. This is the reason for the voice command button – to notify the earset agent application 320 to expect a command.

Operation of the earset communication system in accordance with a preferred embodiment will now be described. As shown in the flow diagram of Figure 9, the user at step 120 may depress the command button 110 on the earset 10 and, after receiving a ready prompt at step 130 from the microprocessor-based appliance 30, the user may speak a command at step 140, such as "Call Mr. Williams," or "Open Microsoft Outlook," or "Close the kitchen blinds," or "What is the temperature outside?" Once the microprocessor-based appliance 30 has received the voice command at step 160 and confirmed the command at step 180 or 190, then the system software initiates the appropriate action at step 210.

If the earset system is making a call, the connection to the network 80, shown in Figure 5A, preferably is muted while the command is issued and being responded to so the remote party does not hear the command. If the command was not recognized at step 170 or at step 220, then the user may again be prompted or asked to start over at step 110 or at step 140.

Preferably, the system utilizes the Lernout & Hauspie speech recognition engine model # ASR 1600/M, which requires no voice training, no names or numbers to enter (assuming that the user already has names and numbers recorded in a contact management/address book system like Microsoft Outlook, Lotus Notes, Windows address book, etc.), and no learning curve to go through. One skilled in the art may readily adapt any appropriate commercially available speech recognition engine. The voice recognition engine will also preferably support multiple or alternative languages for example, English, Spanish, German, Chinese, French, Japanese to name a few. The system may use the names that already exist in the user's contact file, through a dynamic interface to Microsoft Outlook, ACT, Lotus Organizer, and similar products.

The software that operates the system may be an application based on the Microsoft Windows 98 or Windows 2000 operating system (or any subsequent release) and will preferably comply with the "Designed for Microsoft Windows" Logo program, to which those interested may refer. For the preferred embodiment of Figure 7A, the system preferably includes an open hardware platform for multimedia playback and recording as well as button press events.

For the alternative embodiment of Figure 7B, the system preferably includes an open hardware platform for telephony utilizing Microsoft's Telephony API standard. This allows other third party software applications to operate the required system hardware. The system software application uses TAPI 2.0 specification to communicate with the system. The system may also use the TAPI 3.0 specification when available or future versions as they become available.

Microsoft provides support for universal serial bus (USB) 21 using the Microsoft Win32 Driver Model (WDM). Hardware vendors who implement USB solutions for drivers can use the drivers provided by Microsoft or can create minidrivers to exploit any additional unique hardware features. Features requiring a driver that are beyond the functionality of the basic USB audio driver include audio channeling, earset and base station control signaling, telephony control, and the voice command button feature. The base station 20 preferably is a "Plug and Play" device as defined by the Microsoft PC99 (or PC2000) System Design Guide.

The voice agent (VA), also referred to herein as the earset agent application 320, is a speech-based interface agent used to interact with the hardware and other third-party devices and software systems. To accomplish its function, the voice agent utilizes program logic, a speech recognition engine, pre-recorded voice files, and text to speech synthesis where necessary. The VA may use dedicated hardware or other TAPI compliant telephony devices

for its audio I/O and telephony control. In addition, third party hardware and software systems like Savoy's CyberHouse, IBM's ViaVoice and various home automation devices can also be controlled through the VA.

As shown in figure 10C, the process for initiating the voice agent is by pressing the voice command button 110 on the earset 10 or base station 20 (speakerphone) to activate the VA at step 120. This activation plays the ready prompt at step 130 of figure 10C through the earset speaker 52 and places the VA in a listening state. The period of time for placing the VA in a listening state is a system configurable option: for example, 2 seconds. If no speech is detected, the system will revert to its previous state. Further details on the activation of the voice agent and the ready prompt are provided below with reference to the description of the various use cases.

The ready prompt may consist of a user recorded audio stream (WAV file), a pre-selected application-offered audio stream, or a simple combination of tones. The ready prompt will be an application configurable variable. For example, the ready prompt may consist of:

"Yes Steve?"

For purposes of this explanation, all voice command dialogs will assume the voice command button 110 has been pressed and the ready prompt has been played.

In another preferred embodiment, the system is capable of answering the phone and asking the remote party their name and who they are calling. The call may then be announced through wired or wireless speakers located strategically around the house or office that are controlled by the microprocessor-based appliance 30 running the software so the residents know who should answer the phone, and who is calling. This feature can also be used for paging and general announcements.

For example the software can screen out telemarketing calls. Many telemarketers use predictive dialers, which are simply computer programs that dial phone numbers and wait for a human to answer the phone. Telemarketing calls by telemarketers using predictive dialers are screened out automatically because their predictive dialer software makes the determination that a person has not answered the telephone and hangs up. The system may also identify the caller thus eliminating the need for Caller ID. Individual speakers in each room can be selected by the user or automatically by the software so that people may be paged and people may join a conversation. In a home application, the system may announce when vehicles have pulled into the driveway, when any door has been opened, when there are visitors at the front door and when mail has arrived.

Telephony Service Provider (TSP)

A telephony service provider is a dynamic-link library (DLL) that supports communications over a telephone network through a set of exported service functions. The service provider responds to a telephony request, sent to it by the TAPI, by carrying out the low-level tasks necessary to communicate over the telephone network. In this way, the service provider, in conjunction with TAPI, shields applications from the service and technology dependent details of the telephone network communication.

Each service provider is responsible for responding to telephony requests from TAPI to control lines and telephone devices. A service provider is also responsible for controlling and assessing the information exchanged over a call. To manage this information (called the media stream), the service provider must provide additional capabilities or functions. The System TSP may optionally have configuration options to interface with PBX commands. These configuration options define what the flash, park, transfer, conference, forward, etc.

commands equate to in terms of hook flash commands. For example, a conference command may consist of "flash *2".

Issuing a Command

The processing of a command issued via the earset device 10 will now be described with reference to the flow chart of figure 9. Figure 9, which is preferably implemented in software, depicts a preferred method for handling a command issued by a user. As shown in Figure 9, and further described below, the method preferably includes the ability to handle recognition errors. It will be recognized upon review of the following that Figure 9 depicts a generalized method for issuing a command. Specific examples of particular commands will be presented separately below. Figures 5A and 5B illustrate the audio signal paths within the earset communication system associated with the general method.described in Figure 9.

As shown in Figure 9, initiation of the processing of a user command begins at step 115, where the initial conditions of the earset communication system are as follows: (1) the microprocessor-based appliance 30 is powered on; 2) the base station 20 is connected to the microprocessor-based appliance 30, such as via a USB port; 3) the base station 20 is powered on; and 4) a voice agent communication software application is running on the microprocessor-based appliance 30.

At step 120, the user presses the command button 110 on the earset 10, shown in Figure 2, which causes the earset 10 to transmit a signal to the microprocessor-based appliance 30, through the base station 20. Upon receipt at the microprocessor-based appliance 30, the signal activates the voice agent. As described above, the voice agent is preferably a speech-based interface agent used to interact with the system hardware and other third party software products, such as Microsoft Outlook, Lotus Notes, Lernout and Hauspie Voice Express, Dragon Dictate (from Dragon Systems), VoIP capable software (Net2Phone,

DialPad, Microsoft NetMeeting), Instant Messaging Products (ICQ, AOL Instant Messenger, Yahoo! Messenger), or any other voice enabled applications or applications that could benefit from being voice enabled. A suitable, commercially available voice agent is the Arial Voice Agent, offered by ArialPhone LLC of Vernon Hills, Illinois. In response to the signal, the microprocessor-based appliance 30 issues a ready prompt at step 130 to the earset 10 and places the voice agent in a listening state for in a pre-configured manner. In a preferred embodiment, the ready prompt in the application may be configurable in one of many user selectable ways. For example, the ready prompt may be an audio stream containing a message pre-recorded by the user, a generic pre-selected audio stream offered by the application software, or a simple earcon signal characterized by short bytes or tones that are associated with a specific event.

In response to the ready prompt, the user may issue a verbal command at step 140. At step 150 the system determines whether the user spoke. If the user does speak, then the method proceeds to step 160, where voice recognition processing is performed on the command. If the system detects silence, i.e. the user does not speak, then the method proceeds to step 152, where the user is re-prompted. In accordance with a preferred embodiment, the number of times that the user may be re-prompted is a configurable option. The preferred number of re-prompts, for usability purposes, is 2 times total – i.e., initial command and 1 re-prompt. More than this tends to frustrate users, however the number of re-prompts is configurable so more tolerant users can set it higher. In this case, the system then determines whether the user has been re-prompted the predetermined number of times at step 156. If the user has not yet been prompted the maximum number of times, then the method returns to step 140 so the user may issue a command. If, on the other hand, the user has been prompted the predetermined number of times, then the method proceeds to step 240, where

the user is informed of the failure to recognize a command and then the system returns to step 115.

Returning to the case where the user does speak, at step 160 the voice recognition processor associated with the voice agent preferably returns recognition confidence level information, which may be used to determine how accurately a phrase, in this case a command, was recognized. The speech recognition processor preferably assigns a confidence level to the spoken command and then sorts the assigned confidence level into one of three recognition quality categories: high confidence (for example, above 90% confidence), low confidence (for example, between 70%-90% confidence), and unrecognizable (for example, below 70% confidence). In the most favorable situation, the confidence in the speech recognition is high and the method proceeds to step 190 where the PC implicitly verifies the issued command and opens a recognizer. An implicit verification is characterized in that the user is not prompted to verbally confirm the command because of the high confidence in recognizing the spoken command. At step 195, the method determines whether the user has cancelled the confirmed command. If so, the method returns to step 130 where the earset 10 plays the ready prompt to let the user know they can restate the command. If on the other hand, the user does not cancel the confirmed command at step 195, then the method proceeds to step 210 where the command is executed.

If the confidence in the speech recognition is, for example, between 70% and 90%, then the confidence is categorized as low at step 160, and the method proceeds to step 180, where the earset agent application 320 sends a command verification prompt to the user. For example, command verification may comprise repeating the command and asking the user to verbally confirm its accuracy. Specifically, the user may hear through the speaker on the earset, "Did you say 'call John Doe'?" At step 200, the method determines whether the user

replies affirmatively to the command verification prompt. If so, then the method proceeds to step 210 and the command is executed. If, however, the user does not reply affirmatively to the command verification prompt at step 200, then the reply is characterized as unrecognizable, and the user is re-prompted, at step 220, for a command. The number of times to re-prompt the user is preferably a configurable option. Silence by the user during the configurable response period may be treated as an unrecognizable response at step 220. If the user has been re-prompted the predetermined number of times without resulting in an affirmative response, then the method proceeds to step 240. If the user has not been re-prompted the predetermined number of times, then the method returns to step 180.

If the spoken command is unrecognizable based, for example, on a less than 70% confidence in recognition of the command at step 160, then the method proceeds to step 170, where the user is re-prompted, preferably repeatedly for a predetermined number of times. Once the user has been re-prompted the predetermined number of times, as determined at step 172, without the voice agent receiving a recognizable command, then the user is informed at step 176 of a failure to recognize the command, and the method returns to step 115. As noted above, the number of repetitions is preferably a user configurable option. If the user has not been re-prompted the predetermined number of times, then the method proceeds from step 172 to step 140 and the system awaits the user's command.

Making A Call

Turning now to specific examples of particular commands, a preferred embodiment of the present invention allows the user to place a call using the earset communication system. Figure 10 is a flow chart illustrating the basic course for making a call. Beginning at step 250, the user requests the voice agent. This corresponds to steps 115, 120 and 130 in Figure 9.

Next, at step 260, the user issues a command in a predetermined form to indicate to the communication system the user's desire to place a call. Preferably, the earset agent application 320 recognizes synonyms for commonly used commands. For example, the "call" command may be recognized whether the user says "call", "dial" or "get me." Preferably, the generalized method of Figure 9 is followed in regard to recognition rates and the process in the event that the command is not recognized. The actual command may request that the system call a person at a particular location. For example, the user may use a command, "Call Steve Smith at Work." At the microprocessor-based appliance 30, the voice agent will therefore process the command for recognition of 1) the type of command, such as a call; 2) the person to call; and 3) the location.

Once the command has been recognized, the method proceeds to step 270, where the voice agent looks up the called party's number, such as an IP address or telephone number, at the requested location. Generally, the user's contacts are stored in memory at the microprocessor-based appliance 30. For example, the microprocessor-based appliance 30 may include a software application for storing and accessing contact information. There are numerous software applications that are suitable for this purpose including, for example, Microsoft Outlook, which is available from Microsoft Corp of Redmond, Washington and Lotus Notes, which is available from Lotus Development Corporation of North Reading, Massachusetts. Step 270 and the following steps of Figure 10 correspond to the step 210 of Figure 9.

The voice agent then confirms the command to call the called party at step 280. For example, the voice agent implicitly confirms the users request by stating to the user, "Calling Steve Smith at Work." If the user does not cancel the confirmed command, the method proceeds to step 290, where a call is placed to the called party at the desired location. If

however an explicit confirmation is required, for example where the confidence in the speech recognition of the command is Low Confidence or Unrecognizable, then the method preferably proceeds along the paths of steps 180 or 170, respectively, in Figure 9. Reference may be made to the flow chart in Figure 9 for further detail regarding command confirmation. Again, once the command is confirmed, either implicitly or expressly, the method proceeds to step 290 for execution by placing the call.

Making A Call - Alternative Course 1

The flow chart in Figure 11A shows an alternative method for placing a call using the earset communication system. The method shown in Figure 11A generally follows the method of Figure 10, except that the system checks that the requested location for a particular person being called is valid. Thus steps 250 and 260 are the same in Figure 10, except that the method of Figure 11A requires the user to specify a location for the called party. This method is necessary where a called party has multiple phone numbers designated by a unique location such as home or work. Likewise, steps 280 and 290 are present in both embodiments.

Following the user's issuance of a command to call the called party at a particular location at step 260, the method of Figure 11A proceeds to step 305, where the system determines whether the requested location is valid. Generally, a requested location will be considered valid if the user's contact information includes a number for the called party at the requested location.

If the requested location is valid, then the method proceeds to step 310, where the voice agent determines the called party's number at the requested location. From there, the method proceeds to place the call to the called party at the requested location in accordance with steps 280 and 290, which are described above. If, on the other hand, the requested

location is invalid, then the method proceeds to step 325, where the voice agent informs the user that the location is not valid. For example, the voice agent in step 325 may respond with: "That's not a valid location; you can say [location_1]...[Location_n]," where [location_1]...[Location_n] correspond to the valid locations associated with the called party. Alternatively, in the event where there is no number defined for a requested location, the system may prompt the user to enter one. Since each called party may have numerous numbers corresponding to different locations, for example, home, work, mobile and the like, the system will preferably inform the user of each valid location. Next, the user responds with the desired location at step 335. The method then returns to step 305 in order to determine if the location is valid. Once the location information is determined to be valid at step 305, then the method proceeds with steps 310, 280 and 290 as described above.

Making A Call - Alternative Course 2

The flow chart in Figure 11B shows another alternative embodiment of the method for placing a call using the earset communication system. The initial steps are similar to the initial steps in Figure 11A, except that in Figure 11B the user command at step 260 includes only the called party's name. The method proceeds to step 345, where the system determines whether there is more than one number assigned to the called party's name. If more than one number is assigned to the called party's name, then the method proceeds to step 355, where the voice agent prompts the user for more information, such as by requesting "At which location?"

At step 365, the user will respond to the prompt by speaking the location desired for the called party. The system then determines, as described above with reference to Figure 11B, whether the location specified by the user is valid at step 305 and the method progresses as described with reference to Figure 11B.

Returning to step 345, if the method determines that there is only one number for the called party, then the method proceeds to step 375, where the voice agent determines the

proper number from the user's contact information. The method then proceeds to steps 280 and 290, which are described above, to complete placement of the call to the called party.

Retrieving Schedule Information

A preferred embodiment of the present invention allows the user to retrieve schedule information using the earset communication system. Figure 12 is a flow chart illustrating the basic course for retrieving schedule information . Beginning at step 250, the user requests the voice agent. This corresponds to steps 115, 120 and 130 in Figure 9.

Next, at step 262, the user issues a command in a predetermined form to indicate to the communication system the user's desire to retrieve schedule information. Preferably, the generalized method of Figure 9 is followed with regard to recognition rates and the process in the event that the command is not recognized. The actual command may request for a description of the user's schedule. For example, the user may use a command, "What is my schedule today?" At the microprocessor-based appliance 30, the voice agent will therefore process the command for recognition of the user's schedule.

The voice agent then confirms the command to retrieve schedule information for today at step 282. For example, the voice agent implicitly confirms the users request by stating to the user, "Retrieving schedule information for today." If the user does not cancel the confirmed command, the method proceeds to step 288, where the schedule information is retrieved. If however an explicit confirmation is required, for example where the confidence in the speech recognition of the command is Low Confidence or Unrecognizable, then the method preferably proceeds along the paths of steps 180 or 170, respectively, in Figure 9. Reference may be made to the flow chart in Figure 9 for further detail regarding command confirmation. Again, once the command is confirmed, either implicitly or expressly, the

method proceeds to step 288 for retrieving the schedule. Note that the user can interrupt and issue commands such as "next item", "previous item", "next day", "cancel", etc.

Once the command has been recognized, the method proceeds to step 288, where the voice agent looks up the user's schedule. Generally, the user's schedule is stored in memory at the microprocessor-based appliance 30. For example, the microprocessor-based appliance 30 may include a software application for storing schedule information. There are numerous software applications that are suitable for this purpose including, for example, Microsoft Outlook, which is available from Microsoft Corp of Redmond, Washington and Lotus Notes, which is available from Lotus Development Corporation of North Reading, Massachusetts. Finally, at step 292, the voice agent reads or plays the requested schedule information to the user based on the user's previous command.

Control of Home Entertainment and Home Automation

In another preferred embodiment, the earset communicator system functions with existing home control and home entertainment applications that rely heavily on devices such as remote controls and PC-based software interfaces to control various home functions. Implementing voice-based command and control of home functions using the earset communicator system greatly improve convenience and simplicity to the control of the home. Existing IR remote control units are limited to line-of-sight operation and require multiple button sequences to be learned and pressed for most operations. The earset communication system works from anywhere in the home and can respond to natural language commands, such as "Put on ESPN". Functions that may be under voice control include: Television, Digital Music, DVD, Gaming, Lighting, HVAC (Heating Ventilation Air Conditioning), Motorized Blinds and the like.

Figure 13 is a flow chart illustrating the steps for the control of home entertainment functions. Figure 14 is a flow chart illustrating the steps for the control of home automation functions. The steps referenced below refer to software flow diagrams of both figures 13 and 14. Beginning at step 250, the user requests the voice agent. This corresponds to steps 115, 120 and 130 in Figure 9.

Next, at step 264 in Figure 13, the user issues a command in a predetermined form to indicate to the communication system the user's desire to control a home entertainment device. For example, the command may request that the TV be tuned to a particular channel as shown in Figure 13. Preferably, the generalized method of Figure 9 is followed with regard to recognition rates and the process in the event that the command is not recognized. The voice agent then implicitly confirms the command to control or adjust the home entertainment device at step 284. For example, the voice agent implicitly confirms the users request by stating to the user, "Tuning TV to ESPN" for the control of the TV. If the user does not cancel the confirmed command, the method proceeds to step 294, where upon execution of the command, the home entertainment device is controlled in the manner commanded by the user. Reference may be made to the flow chart in Figure 9 for further detail regarding command confirmation.

The generalized flow chart shown in Figure 14 illustrates the software flow for adjustment or control of a home automation function. The format of the generalized home automation command may be <adjustment or control> of the <home automation function> where the item in the <field> indicated is a command variable. The user issues such a command at step 266 in Figure 14. The actual home automation function may be, for example, to lower the kitchen blinds. The voice agent software running on the microprocessor-based appliance 30 will therefore process the command for recognition of the

command and for identification of the appliance to be controlled. At step 286 the voice agent confirms the command. If the user does not cancel the confirmed command, the method proceeds to step 296, where upon execution of the command, the appliance is controlled in the manner commanded by the user. If however an explicit confirmation is required, for example where the confidence in the speech recognition of the command is Low Confidence or Unrecognizable, then the method preferably proceeds along the paths of steps 180 or 170, respectively, in Figure 9. Reference may be made to the flow chart in Figure 9 for further detail regarding command confirmation. Again, once the command is confirmed, either implicitly or expressly, and the method proceeds to step 296 for execution of the command. Once the command has been recognized, the method proceeds to step 280, where the appliance is adjusted in the desired manner.

As an alternative to the foregoing, the earset communication system may be embedded with a home appliance or home entertainment device, provided that the appliance or device includes a read-only memory (ROM) structure, a random access memory (RAM) structure, associated data and address buses, and a port for coupling the appliance or device to the base station 20. One skilled in the art will readily adapt control of appliances to other home automated appliances such as home controller links with actuators for curtains, blinds, lights, garage door openers, video cameras, TVs and intercoms to name just a few possible home appliances that may be valid appliances in the <home automation function> field. The <adjustment or control> field for example could be an on/off operation, up/down volume, open/close, or other change mode function.

We claim:

1. A method of communicating with a telecommunications system, through a microprocessor-based appliance having a memory structure, over a wireless link, comprising the steps of:

receiving an audio signal at a receiver;

transmitting an audio signal at a transmitter;

processing the transmitted audio signal to recognize a command; and

controlling the microprocessor-based appliance to effect a desired mode of communication on the telecommunications system based on the command.

2. The method of controlling the microprocessor-based appliance as claimed in claim 1, wherein the mode of communication is a VoN call.

3. The method of controlling the microprocessor-based appliance as claimed in claim 1, wherein the mode of communication is a telephone call.

4. A wireless communication system communicating with a telecommunications network comprising:

an earset having a transmitter and receiver;

a base station having a transmitter and receiver in communication with the earset;

a microprocessor-based appliance, having a memory structure, connected to the base station; and

a speech processing program in the memory and executable by the appliance, said speech processing program associated with the control of a communication system.

5. The apparatus as claimed in claim 4, wherein the earset comprises a single control button that when depressed alerts the speech processing program that an immediately following audio stream is to be interpreted as a command.

6. The apparatus as claimed in claim 4, where the connection between the base station and the microprocessor-based appliance is through a universal serial bus port.

7. The apparatus as claimed in claim 4, wherein the microprocessor-based appliance further comprises a network interface.

8. The apparatus as claimed in claim 4 where the earset further, comprises:

an antenna connected to the transmitter ;

a microphone coupled to the transmitter;

a receiver connected to the antenna; and

a speaker coupled to the receiver.

9. A method of communicating in a user-communication interface over a wireless link between an earset and a base station comprising the steps of:

transmitting an audio signal into a modulated carrier;

receiving the modulated carrier to produce an audio signal;

sending a command signal from the earset to the basestation;

performing voice recognition on the audio signal after receiving the command signal; and

sending control signals based on a voice command.

10. An earset communication system that provides a user with a hands-free interface for communication and control functions, comprising:

a lightweight wireless earset having a speaker, a microphone and a radio transceiver coupled to the speaker and microphone;

a microprocessor-based appliance having a network interface and software for communication and control of a plurality of subsystems based on speech recognition capabilities; and

a base station having a first interface for communicating with the earset over a radio link and a second interface for communicating with the microprocessor-based appliance;

wherein the user speaks a command into the microphone and the microprocessor-based appliance drives a selected subsystem to execute the command.

11. The system of claim 10, wherein the command causes the microprocessor-based appliance to initiate a VoN call with a predetermined remote party.

12. The system of claim 10, wherein the command causes the microprocessor-based appliance to initiate a PSTN telephone call.

13. The system of claim 10, wherein the command causes the microprocessor-based appliance to adjust a home entertainment appliance.

14. The system of claim 10, wherein the command causes the microprocessor-based appliance to accept voice dictation.

15. The system of claim 10, wherein the command causes the microprocessor-based appliance to adjust an automated home appliance.

16. A method for a microprocessor-based appliance to communicate with an earset, the method comprising:

receiving a voice signal from the earset, the voice signal comprising a command;

recognizing the command in the voice signal; and

performing a function in response to recognizing the command.

17. The method of claim 16, wherein the microprocessor-based appliance is connected to a communications network, the method further comprising sending a communications signal to the communications network.

18. The method of claim 16, wherein the function is a VoN call.

19. The method of claim 16, wherein the function is a PSTN call.

20. The method of claim 16, wherein the function is adjustment of a home appliance.

21. The method of claim 16, wherein the function is voice dictation.

22. The method of claim 16, wherein the function is adjustment of a home entertainment appliance.

23. A method for a base station to communicatively couple an earset with a microprocessor-based appliance, the method comprising:

receiving a voice signal from the earset, the voice signal comprising representation of a command;

digitizing the voice signal into a digitized signal; and

sending the digitized signal to the microprocessor-based appliance.

24. The method of claim 23, wherein digitizing the signal comprises:

digitizing the voice signal into an intermediate digitized signal;

converting the intermediate digitized signal into an analog signal; and

digitizing the analog signal into the digitized signal.

25. The method of claim 23, wherein receiving the voice signal from the earset comprises receiving a 900 MHz digital spread spectrum signal.

26. The method of claim 23, wherein sending the digitized signal to the microprocessor-based appliance comprises sending the digitized signal to a Universal Serial Bus port.

27. A method of sending a command from an earset to a microprocessor-based appliance, the earset having a microphone and a command button, the method comprising:

activating the command button on the earset;

receiving a voice signal from the microphone, the voice signal being representation of a command; and

sending the voice signal to the microprocessor-based device.

28. The method of claim 27 further comprising (i) prompting the microprocessor-based device for receipt of the command; and (ii) receiving a ready prompt from the microprocessor-based appliance.

29. The method of claim 27, wherein activating the command-button comprises depressing the command button.

30. The method of claim 27, wherein receiving the voice signal further comprises performing noise cancellation on the voice signal.

31. A microprocessor-based appliance for communicating with an earset, the microprocessor-based appliance comprising:

a processor;

a memory;

computer instructions stored in the memory and executable by the processor for:

recognizing a command in a voice signal received from the earset, the voice signal comprising representation of the command; and

performing a function on the microprocessor-based appliance in response to recognizing the command.

32. The microprocessor-based appliance of claim 31 further comprising a communications interface for connecting the microprocessor-based appliance to a communications network.

33. The microprocessor-based appliance of claim 32, wherein the communications interface is selected from the group consisting of a PSTN interface, a network interface, a Universal Serial Bus port, and a radio link.

34. The microprocessor-based appliance of claim 31, wherein a radio link communicatively couples the earset with the microprocessor-based appliance.

35. The microprocessor-based appliance of claim 31, wherein the radio link is a 900 MHz digital spread spectrum transceiver.

36. The microprocessor-based appliance of claim 31, wherein the function is a VoN call.

37. The microprocessor-based appliance of claim 31, wherein the function is a PSTN call.

38. The microprocessor-based appliance of claim 31, wherein the function is adjustment of a home appliance.

39. The microprocessor-based appliance of claim 31, wherein the function is voice dictation.

40. The microprocessor-based appliance of claim 31, wherein the function is adjustment of a home entertainment appliance.

41. The microprocessor-based appliance of claim 31, wherein a voice agent facilitates performing the function.

42. A base station for communicatively coupling an earset with a microprocessor-based appliance, the base station comprising:

at least one communications interface for (i) receiving a voice signal from the earset, the voice signal comprising a representation of a command; (ii) sending a digitized signal to the microprocessor-based appliance; and

circuitry for digitizing the voice signal received from the earset into a digitized signal.

43. The base station of claim 42 further comprising circuitry for:

(i) digitizing the voice signal into an intermediate digitized signal;

(ii) converting the intermediate digitized signal into an analog signal; and

(iii) digitizing the analog signal into the digitized signal.

44. A base station for communicatively coupling an earset to a microprocessor-based appliance, the base station comprising:

a processor;

a memory;

at least one communications interface for (i) receiving a voice signal from the earset; and (ii) sending a digitized signal to the microprocessor-based appliance, and;

computer instructions stored in the memory and executable by the processor for digitizing the signal received from the earset into the digitized signal.

45. The base station of claim 44 further comprising computer instructions stored in the memory and executable by the processor for:

digitizing the voice signal into an intermediate digitized signal;

converting the intermediate digitized signal into an analog signal; and

digitizing the analog signal into the digitized signal.

46. The base station of claim 44, further comprising mating contacts for charging a battery mounted in the earset.

47. The base station of claim 44, wherein the at least one communications interface for receiving the signal from the earset comprises a 900 MHz spread spectrum transceiver.

48. The base station of claim 44, wherein the at least one communications interface for sending the digitized signal to the microprocessor-based appliance comprises a Universal Serial Bus port.

49. An earset for communicating with a microprocessor-based appliance, the earset comprising:

a command button for prompting the microprocessor-based appliance to receive a command;

a speaker for generating audible sound received from the microprocessor-based appliance;

a microphone for receiving vocal sound, the vocal sound being a command; and

a communications interface for sending a signal to the microprocessor-based appliance, the signal comprising a representation of the vocal sound.

50. The earset of claim 49, further comprising an audio transducer for generating a notice sound.

51. The earset of claim 49 further comprising a communications interface for receiving the audible sound from the microprocessor-based appliance, the audible sound being a ready prompt.

52. The earset of claim 49 further comprising a battery.

53. The earset of claim 49 further comprising mating contacts for charging the battery.

54. The earset of claim 49 wherein the earset comprises a processor and a memory, the earset further comprising computer instructions stored in memory and executable by a microprocessor for performing noise cancellation on the vocal sound.

55. The earset of claim 49, wherein the communications interface for sending a signal to the microprocessor-based appliance comprises a 900 MHz digital spread spectrum transceiver.

56. The earset of claim 49, wherein the communications interface for receiving audible sound from the microprocessor-based appliance comprises a Universal Serial Bus port.

57. The earset of claim 49 further comprising a jack for connecting a separate speaker.

58. The earset of claim 49 further comprising a jack for connecting a separate microphone.

FIG. 1

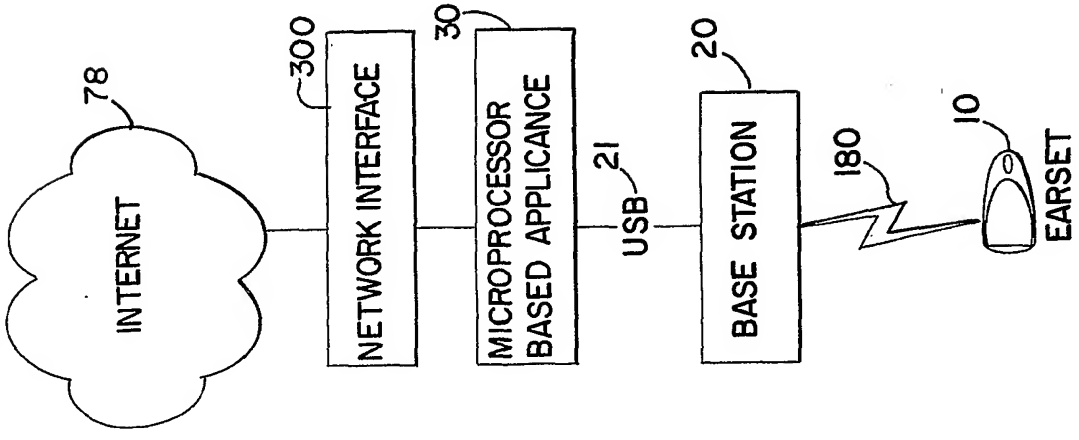


FIG. 2

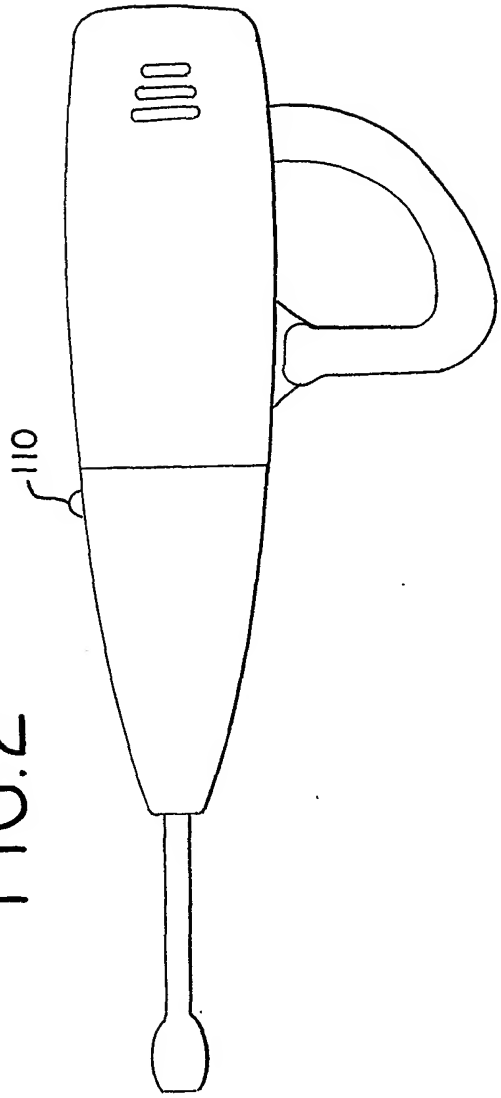


FIG. 3

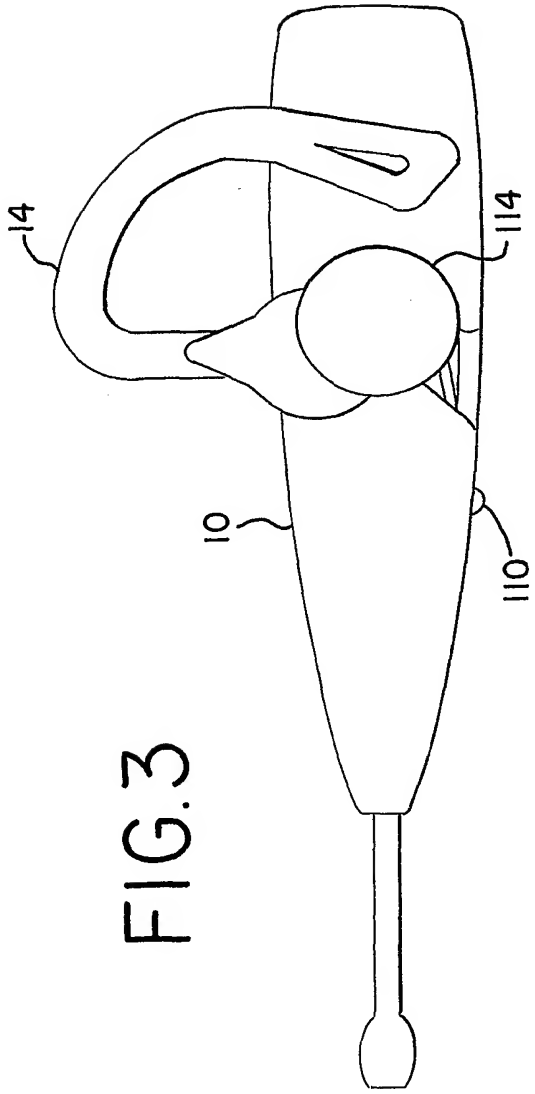


FIG.4A

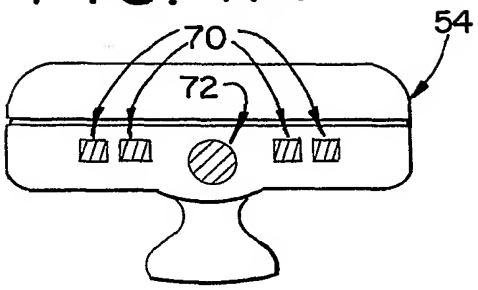


FIG.4B

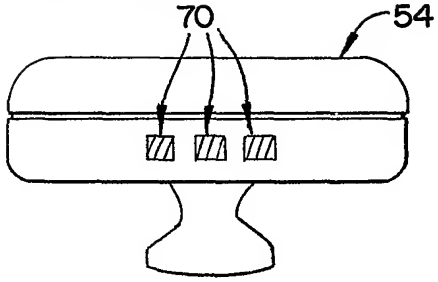


FIG.5A

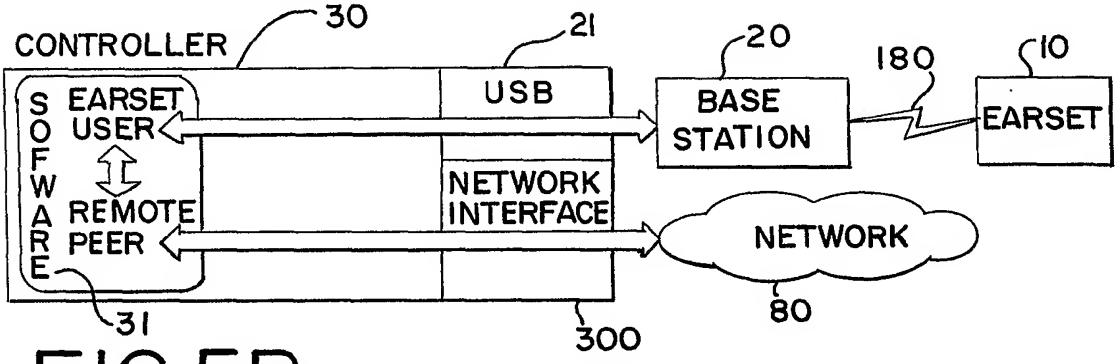


FIG.5B

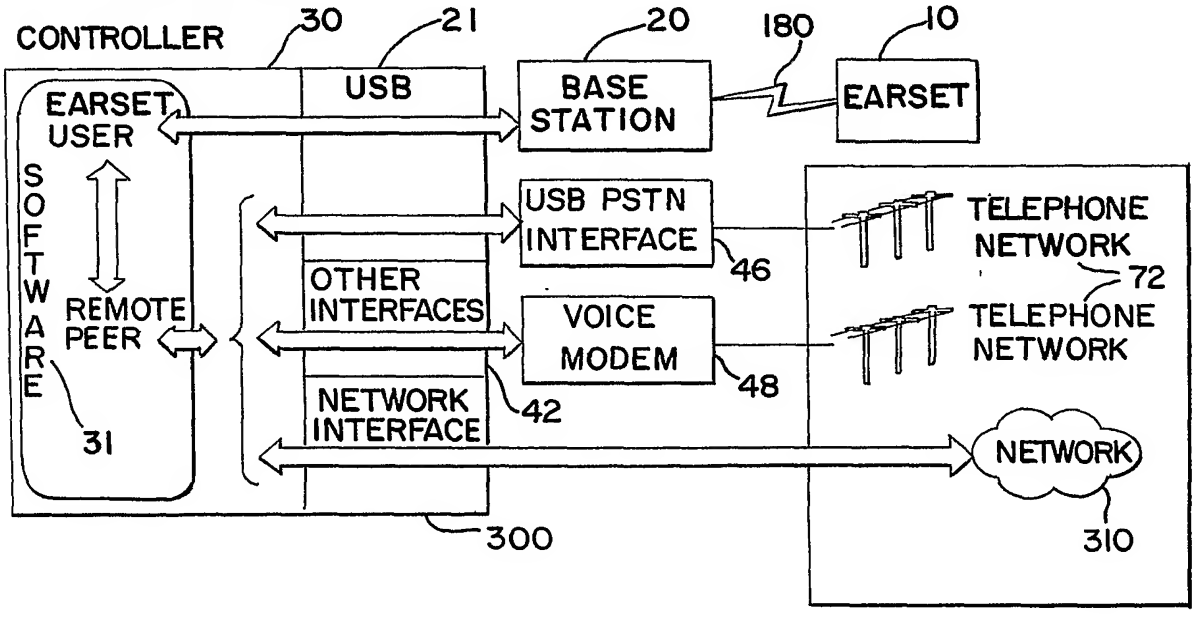


FIG.6

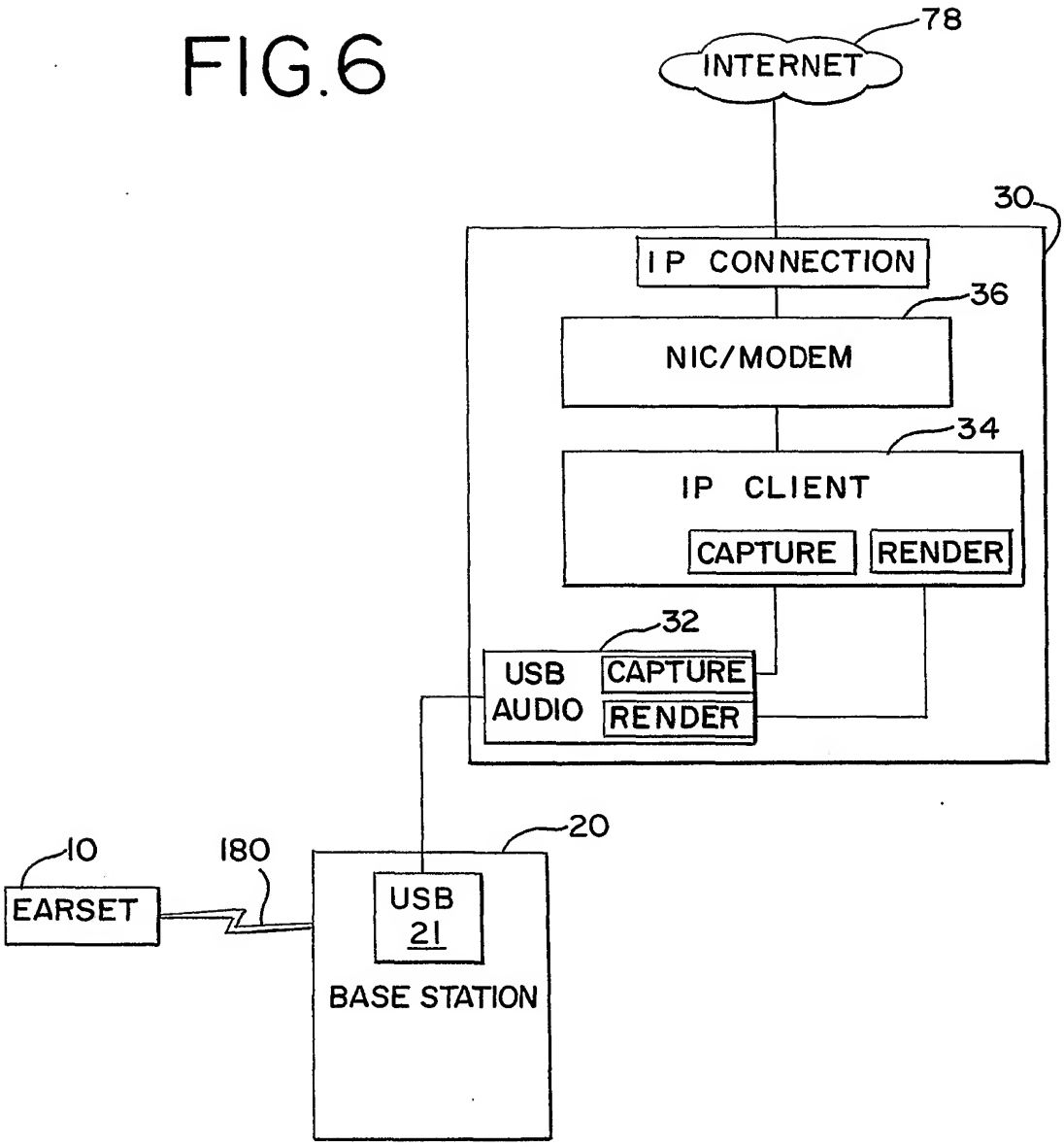

78
INTERNET
30
IP CONNECTION
36
NIC/MODEM
34
IP CLIENT
CAPTURE
RENDER
32
USB AUDIO
CAPTURE
RENDER
20
10
180
EARSET
USB 21
BASE STATION

FIG. 7A

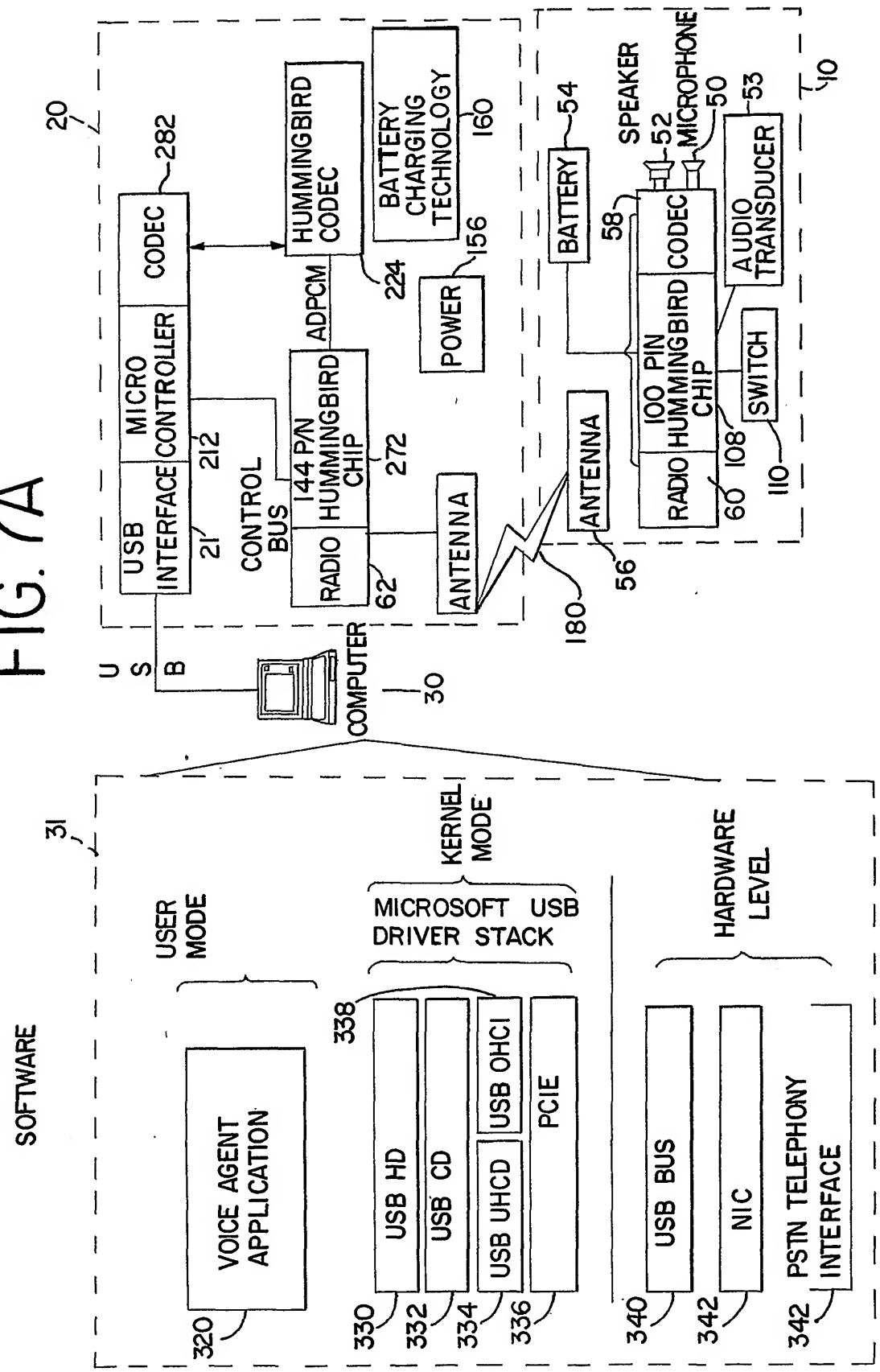

SOFTWARE
31
USER MODE
KERNEL MODE
MICROSOFT USB DRIVER STACK
HARDWARE LEVEL
VOICE AGENT APPLICATION
320
338
USB HD
USB CD
USB UHCD
USB OHCI
PCIE
330
332
334
336
USB BUS
NIC
PSTN TELEPHONY INTERFACE
340
342
342
USB
COMPUTER
30
20
USB INTERFACE
21
MICRO CONTROLLER
212
CODEC
282
CONTROL BUS
144 P/N HUMMINGBIRD CHIP
272
RADIO
62
ADPCM
HUMMINGBIRD CODEC
224
ANTENNA
180
BATTERY CHARGING TECHNOLOGY
160
POWER
156
BATTERY
54
ANTENNA
56
10
SPEAKER
52
MICROPHONE
50
AUDIO TRANSDUCER
53
100 PIN HUMMINGBIRD CHIP
58
RADIO
60
CODEC
SWITCH
108
110

FIG. 7B

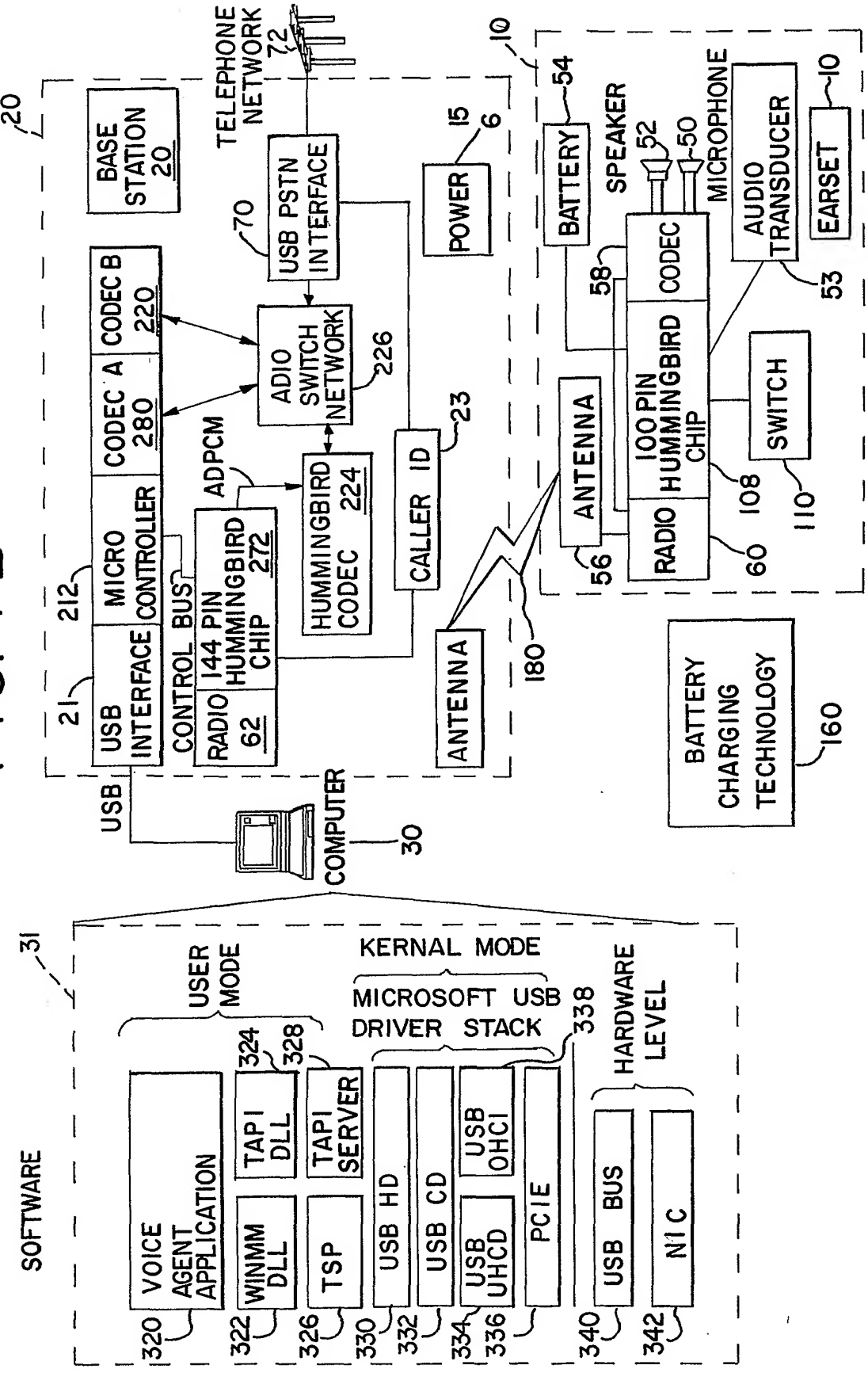

20
BASE STATION 20
TELEPHONE NETWORK 72
USB PSTN INTERFACE 70
POWER 15 6
10
BATTERY 54
SPEAKER 52
MICROPHONE 50
AUDIO TRANSDUCER 53
EARSET 10
CODEC B 220
CODEC A 280
CODEC 58
ADIO SWITCH NETWORK 226
ADPCM
HUMMINGBIRD CODEC 224
CALLER ID 23
ANTENNA 56
100 PIN HUMMINGBIRD CHIP
SWITCH 110
108
RADIO 60
MICRO CONTROLLER 212
144 PIN HUMMINGBIRD CHIP 272
CONTROL BUS
USB INTERFACE 21
RADIO 62
ANTENNA 180
USB
COMPUTER 30
BATTERY CHARGING TECHNOLOGY 160
31
SOFTWARE
USER MODE
KERNAL MODE
MICROSOFT USB DRIVER STACK 338
HARDWARE LEVEL
VOICE AGENT APPLICATION 320
TAPI DLL 324
TAPI SERVER 328
WINMM DLL 322
TSP 326
USB HD 330
USB CD 332
USB UHCI 334
USB OHCI 336
PCIE
USB BUS 340
NIC 342

FIG. 8A

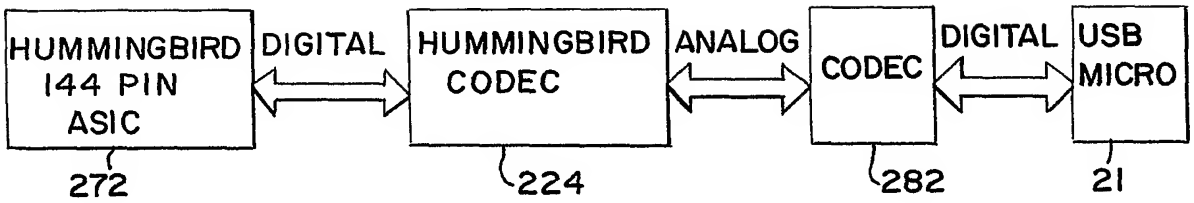

HUMMINGBIRD
144 PIN
ASIC
272
DIGITAL
HUMMINGBIRD
CODEC
224
ANALOG
CODEC
282
DIGITAL
USB
MICRO
21

FIG. 8B

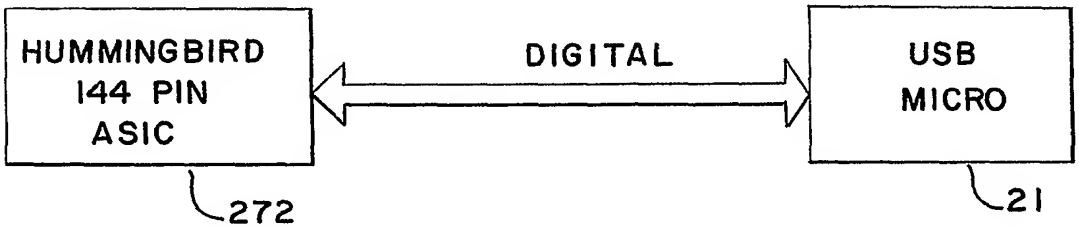

HUMMINGBIRD
144 PIN
ASIC
272
DIGITAL
USB
MICRO
21

FIG. 9

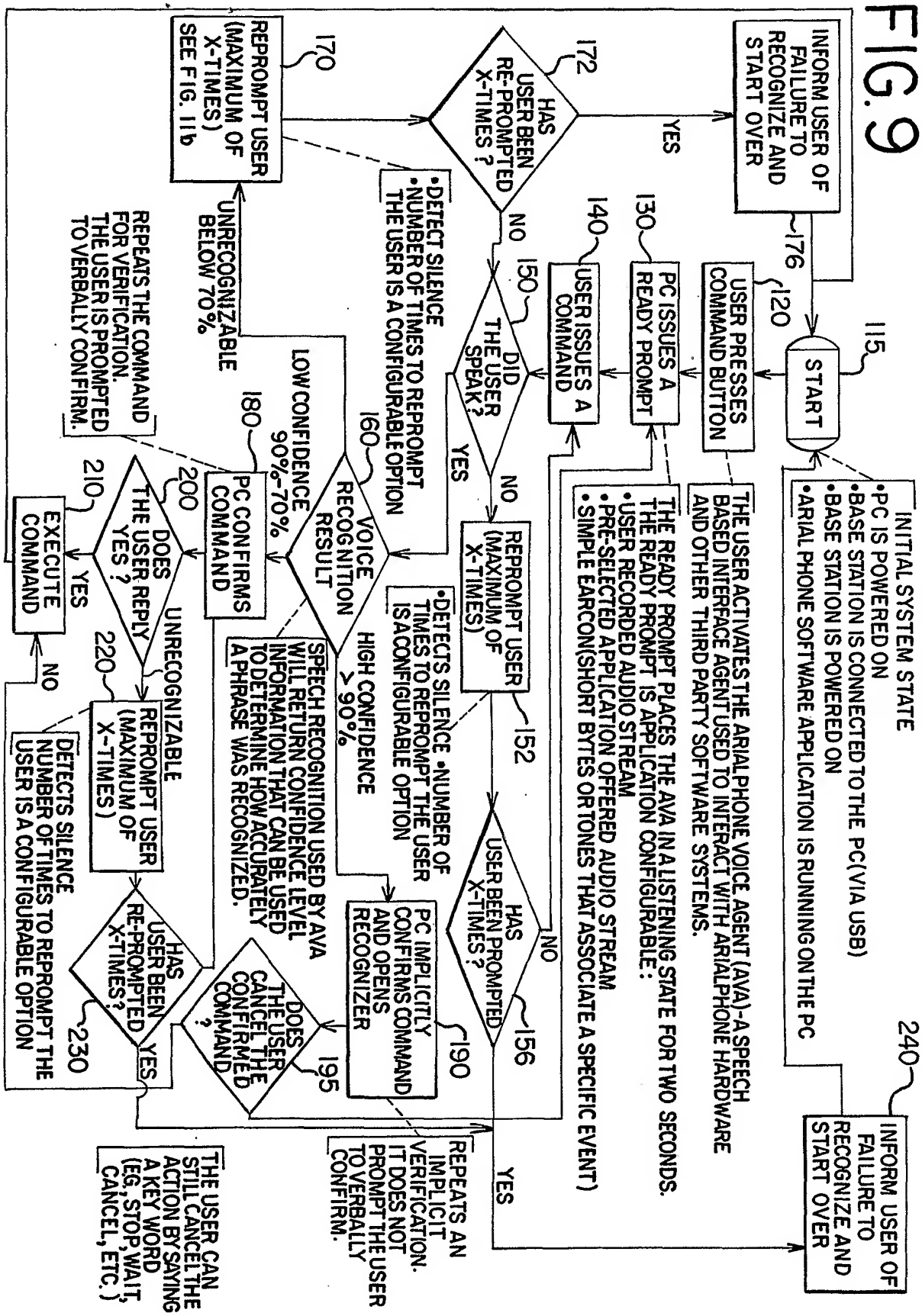

INITIAL SYSTEM STATE
•PC IS POWERED ON
•BASE STATION IS CONNECTED TO THE PC(VIA USB)
•BASE STATION IS POWERED ON
•ARIAL PHONE SOFTWARE APPLICATION IS RUNNING ON THE PC
240
INFORM USER OF FAILURE TO RECOGNIZE AND START OVER
115
START
176
120
USER PRESSES COMMAND BUTTON
130
PC ISSUES A READY PROMPT
140
USER ISSUES A COMMAND
THE USER ACTIVATES THE ARIAL PHONE VOICE AGENT (AVA)-A SPEECH BASED INTERFACE AGENT USED TO INTERACT WITH ARIALPHONE HARDWARE AND OTHER THIRD PARTY SOFTWARE SYSTEMS.
THE READY PROMPT PLACES THE AVA IN A LISTENING STATE FOR TWO SECONDS.
•USER RECORDED AUDIO STREAM
•PRE-SELECTED APPLICATION OFFERED AUDIO STREAM
•SIMPLE EARCON(SHORT BYTES OR TONES THAT ASSOCIATE A SPECIFIC EVENT)
150
DID THE USER SPEAK?
NO
YES
152
REPROMPT USER (MAXIMUM OF X-TIMES)
•DETECTS SILENCE •NUMBER OF TIMES TO REPROMPT THE USER IS A CONFIGURABLE OPTION
156
HAS USER BEEN PROMPTED X-TIMES?
NO
YES
190
PC IMPLICITLY CONFIRMS COMMAND AND OPENS RECOGNIZER
REPEATS AN IMPLICIT VERIFICATION. IT DOES NOT PROMPT THE USER TO VERBALLY CONFIRM.
195
DOES THE USER CANCEL THE CONFIRMED COMMAND ?
THE USER CAN STILL CANCEL THE ACTION BY SAYING A KEY WORD (EG., STOP, WAIT, CANCEL, ETC.)
160
VOICE RECOGNITION RESULT
HIGH CONFIDENCE >90%
SPEECH RECOGNITION USED BY AVA WILL RETURN CONFIDENCE LEVEL INFORMATION THAT CAN BE USED TO DETERMINE HOW ACCURATELY A PHRASE WAS RECOGNIZED.
LOW CONFIDENCE 90%-70%
180
PC CONFIRMS COMMAND
200
DOES THE USER REPLY YES ?
YES
210
EXECUTE COMMAND
UNRECOGNIZABLE
220
REPROMPT USER (MAXIMUM OF X-TIMES)
DETECTS SILENCE
NUMBER OF TIMES TO REPROMPT THE USER IS A CONFIGURABLE OPTION
NO
HAS USER BEEN RE-PROMPTED X-TIMES?
YES
230
REPEATS THE COMMAND FOR VERIFICATION. THE USER IS PROMPTED TO VERBALLY CONFIRM.
UNRECOGNIZABLE BELOW 70%
170
REPROMPT USER (MAXIMUM OF X-TIMES) SEE FIG. 11b
172
HAS USER BEEN RE-PROMPTED X-TIMES ?
YES
NO
•DETECT SILENCE
•NUMBER OF TIMES TO REPROMPT THE USER IS A CONFIGURABLE OPTION

FIG. 10

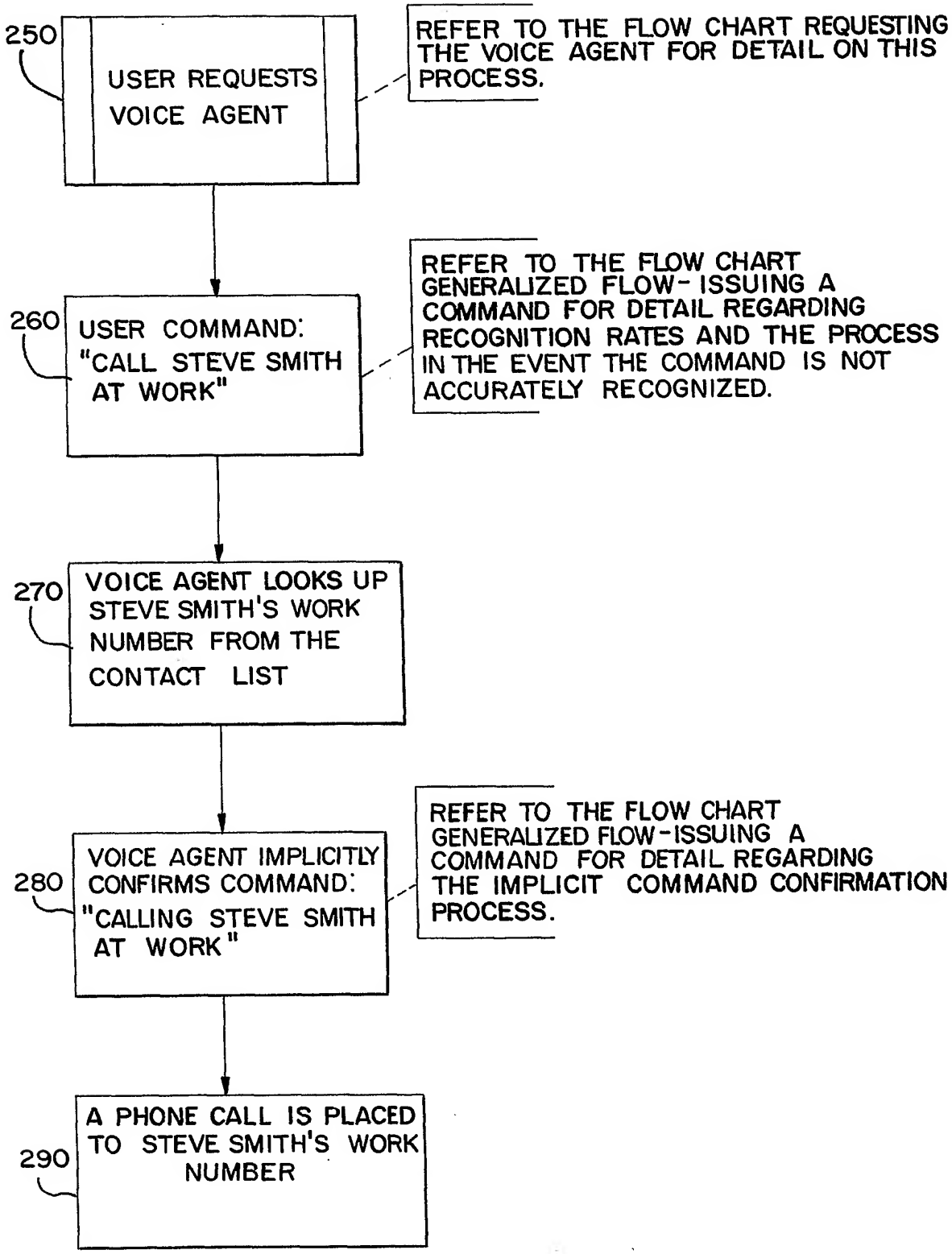

250
USER REQUESTS VOICE AGENT
REFER TO THE FLOW CHART REQUESTING THE VOICE AGENT FOR DETAIL ON THIS PROCESS.
260
USER COMMAND: "CALL STEVE SMITH AT WORK"
REFER TO THE FLOW CHART GENERALIZED FLOW- ISSUING A COMMAND FOR DETAIL REGARDING RECOGNITION RATES AND THE PROCESS IN THE EVENT THE COMMAND IS NOT ACCURATELY RECOGNIZED.
270
VOICE AGENT LOOKS UP STEVE SMITH'S WORK NUMBER FROM THE CONTACT LIST
280
VOICE AGENT IMPLICITLY CONFIRMS COMMAND: "CALLING STEVE SMITH AT WORK"
REFER TO THE FLOW CHART GENERALIZED FLOW-ISSUING A COMMAND FOR DETAIL REGARDING THE IMPLICIT COMMAND CONFIRMATION PROCESS.
290
A PHONE CALL IS PLACED TO STEVE SMITH'S WORK NUMBER

FIG. 11a

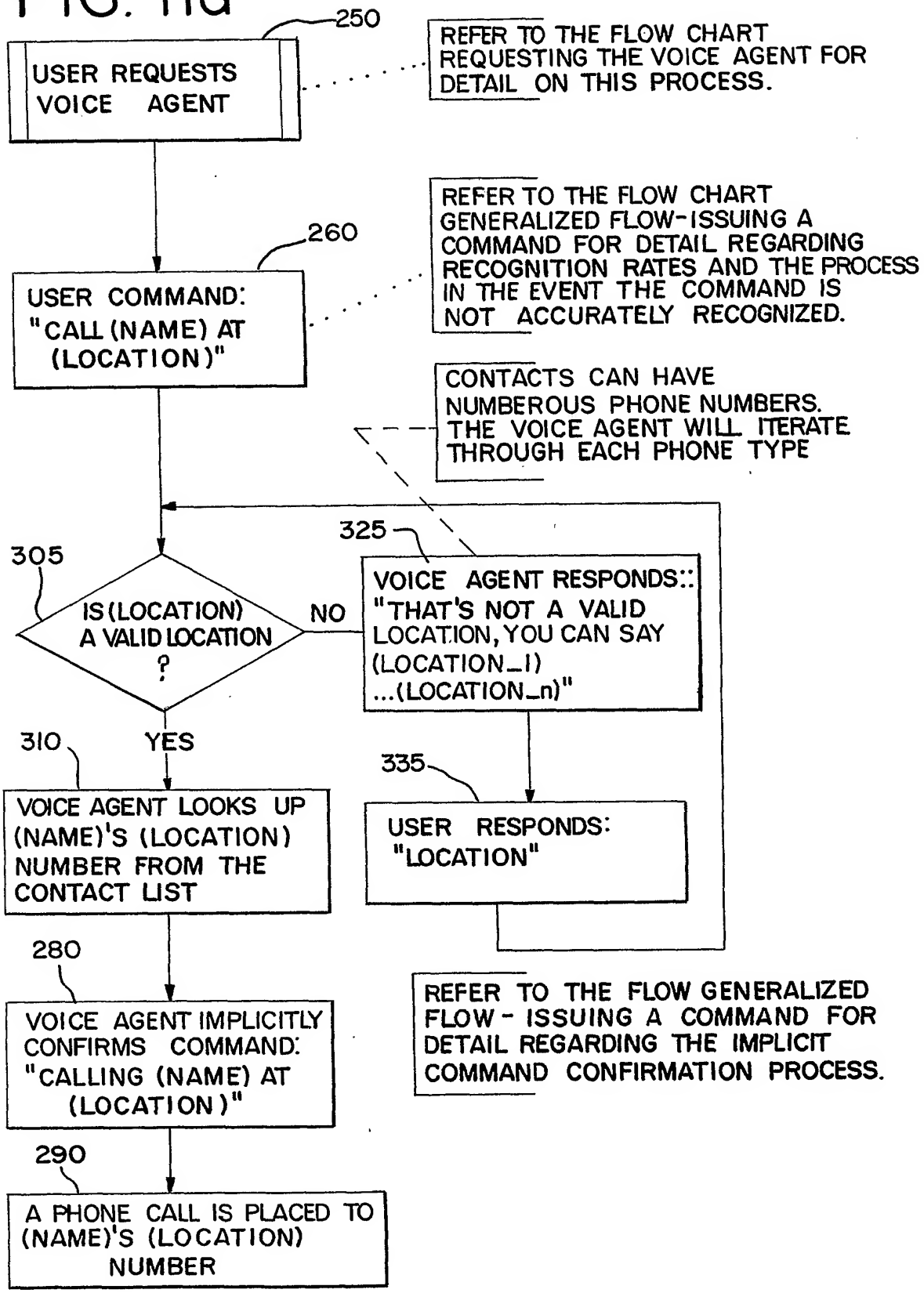

250
USER REQUESTS VOICE AGENT
REFER TO THE FLOW CHART REQUESTING THE VOICE AGENT FOR DETAIL ON THIS PROCESS.
260
USER COMMAND: "CALL (NAME) AT (LOCATION)"
REFER TO THE FLOW CHART GENERALIZED FLOW-ISSUING A COMMAND FOR DETAIL REGARDING RECOGNITION RATES AND THE PROCESS IN THE EVENT THE COMMAND IS NOT ACCURATELY RECOGNIZED.
CONTACTS CAN HAVE NUMBEROUS PHONE NUMBERS. THE VOICE AGENT WILL ITERATE THROUGH EACH PHONE TYPE
305
IS (LOCATION) A VALID LOCATION ?
NO
325
VOICE AGENT RESPONDS:: "THAT'S NOT A VALID LOCATION, YOU CAN SAY (LOCATION_1) ...(LOCATION_n)"
YES
310
VOICE AGENT LOOKS UP (NAME)'S (LOCATION) NUMBER FROM THE CONTACT LIST
335
USER RESPONDS: "LOCATION"
280
VOICE AGENT IMPLICITLY CONFIRMS COMMAND: "CALLING (NAME) AT (LOCATION)"
REFER TO THE FLOW GENERALIZED FLOW - ISSUING A COMMAND FOR DETAIL REGARDING THE IMPLICIT COMMAND CONFIRMATION PROCESS.
290
A PHONE CALL IS PLACED TO (NAME)'S (LOCATION) NUMBER

FIG. 11b

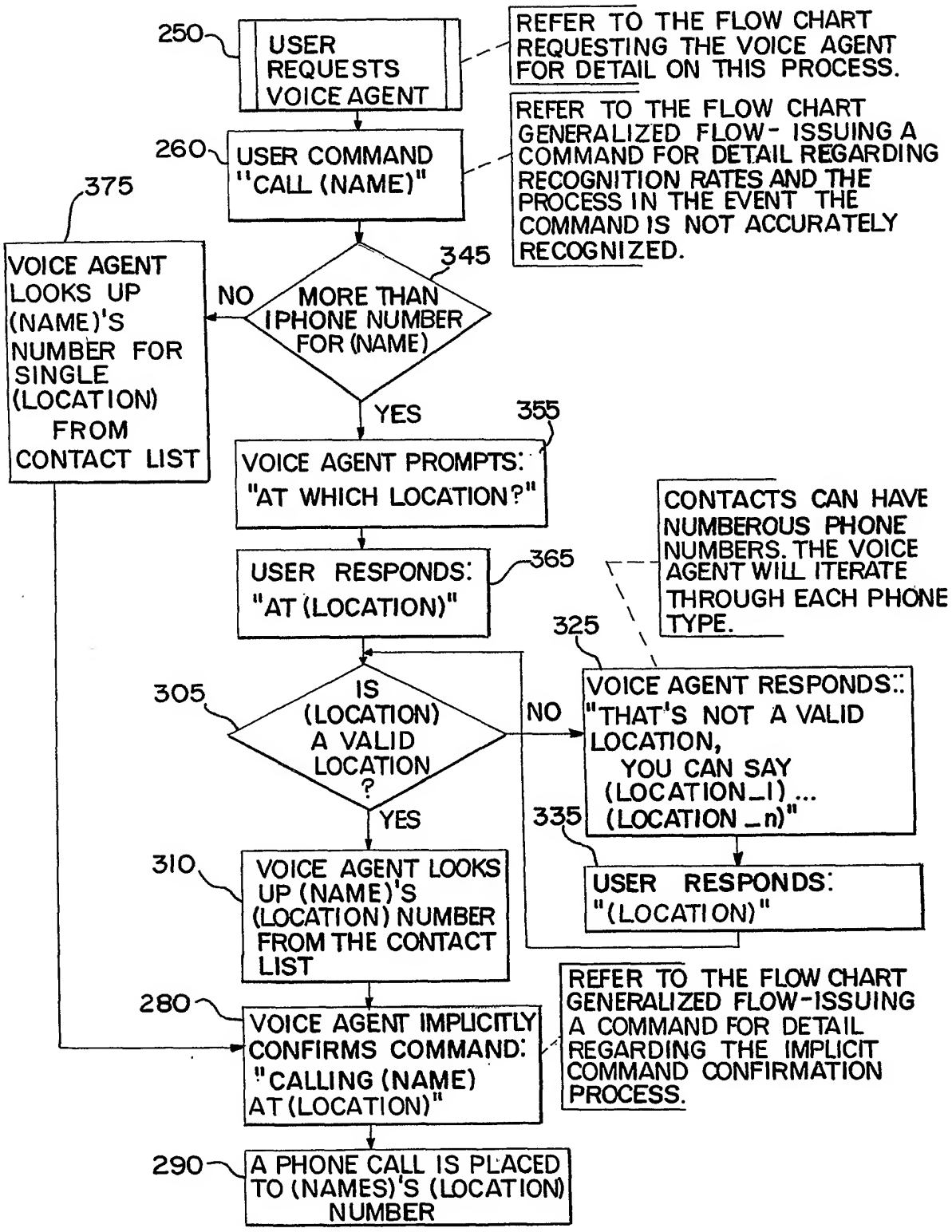

250
USER REQUESTS VOICE AGENT
REFER TO THE FLOW CHART REQUESTING THE VOICE AGENT FOR DETAIL ON THIS PROCESS.
260
USER COMMAND "CALL (NAME)"
REFER TO THE FLOW CHART GENERALIZED FLOW- ISSUING A COMMAND FOR DETAIL REGARDING RECOGNITION RATES AND THE PROCESS IN THE EVENT THE COMMAND IS NOT ACCURATELY RECOGNIZED.
375
VOICE AGENT LOOKS UP (NAME)'S NUMBER FOR SINGLE (LOCATION) FROM CONTACT LIST
345
MORE THAN 1PHONE NUMBER FOR (NAME)
NO
YES
355
VOICE AGENT PROMPTS: "AT WHICH LOCATION?"
CONTACTS CAN HAVE NUMBEROUS PHONE NUMBERS. THE VOICE AGENT WILL ITERATE THROUGH EACH PHONE TYPE.
365
USER RESPONDS: "AT (LOCATION)"
325
305
IS (LOCATION) A VALID LOCATION ?
NO
VOICE AGENT RESPONDS:: "THAT'S NOT A VALID LOCATION, YOU CAN SAY (LOCATION_1) ... (LOCATION _n)"
YES
335
USER RESPONDS: "(LOCATION)"
310
VOICE AGENT LOOKS UP (NAME)'S (LOCATION) NUMBER FROM THE CONTACT LIST
280
VOICE AGENT IMPLICITLY CONFIRMS COMMAND: "CALLING (NAME) AT (LOCATION)"
REFER TO THE FLOW CHART GENERALIZED FLOW-ISSUING A COMMAND FOR DETAIL REGARDING THE IMPLICIT COMMAND CONFIRMATION PROCESS.
290
A PHONE CALL IS PLACED TO (NAMES)'S (LOCATION) NUMBER

FIG. 11C

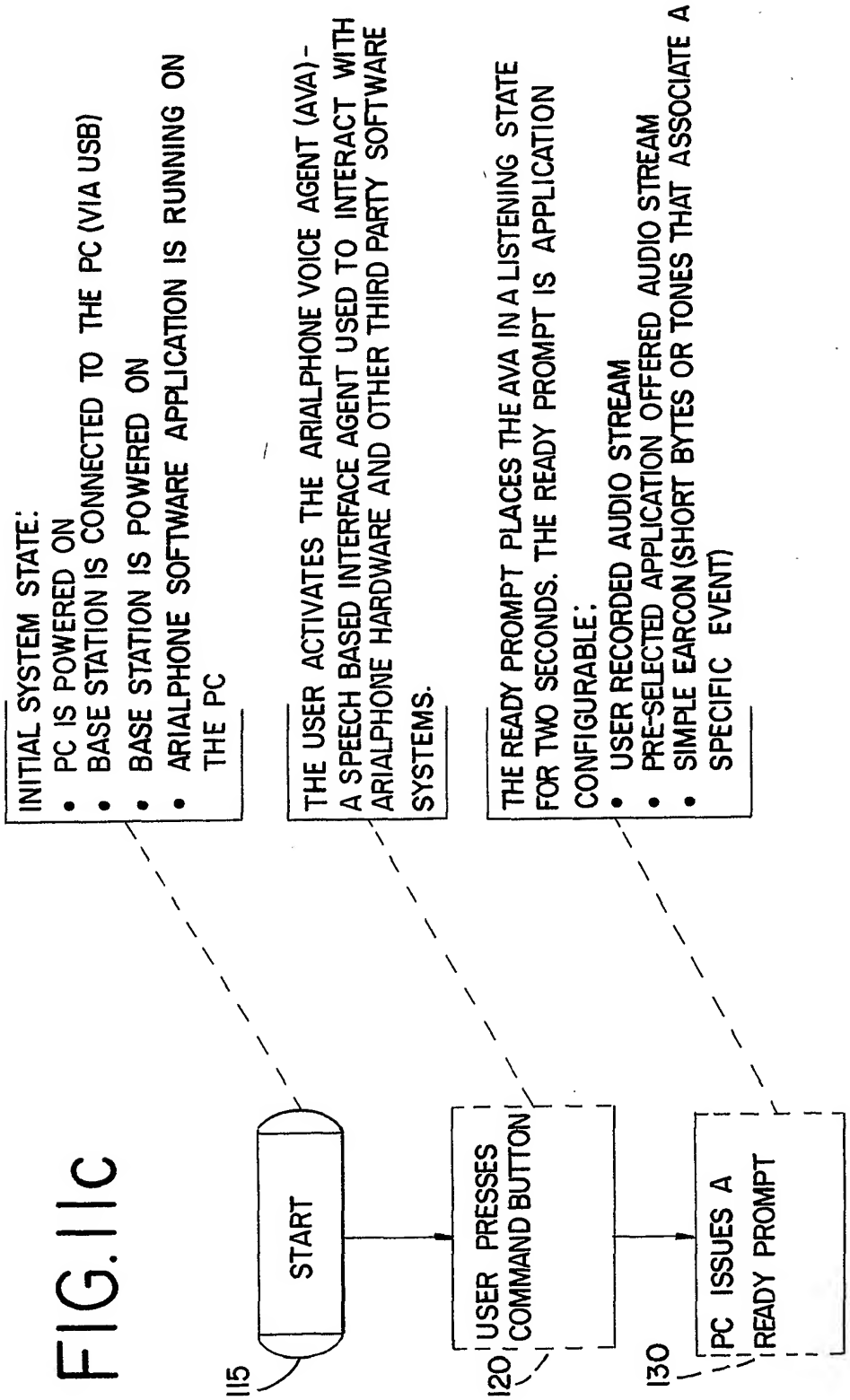

INITIAL SYSTEM STATE:
• PC IS POWERED ON
• BASE STATION IS CONNECTED TO THE PC (VIA USB)
• BASE STATION IS POWERED ON
• ARIALPHONE SOFTWARE APPLICATION IS RUNNING ON THE PC
THE USER ACTIVATES THE ARIALPHONE VOICE AGENT (AVA) - A SPEECH BASED INTERFACE AGENT USED TO INTERACT WITH ARIALPHONE HARDWARE AND OTHER THIRD PARTY SOFTWARE SYSTEMS.
THE READY PROMPT PLACES THE AVA IN A LISTENING STATE FOR TWO SECONDS. THE READY PROMPT IS APPLICATION CONFIGURABLE:
• USER RECORDED AUDIO STREAM
• PRE-SELECTED APPLICATION OFFERED AUDIO STREAM
• SIMPLE EARCON (SHORT BYTES OR TONES THAT ASSOCIATE A SPECIFIC EVENT)
START
115
USER PRESSES COMMAND BUTTON
120
PC ISSUES A READY PROMPT
130

FIG. 12

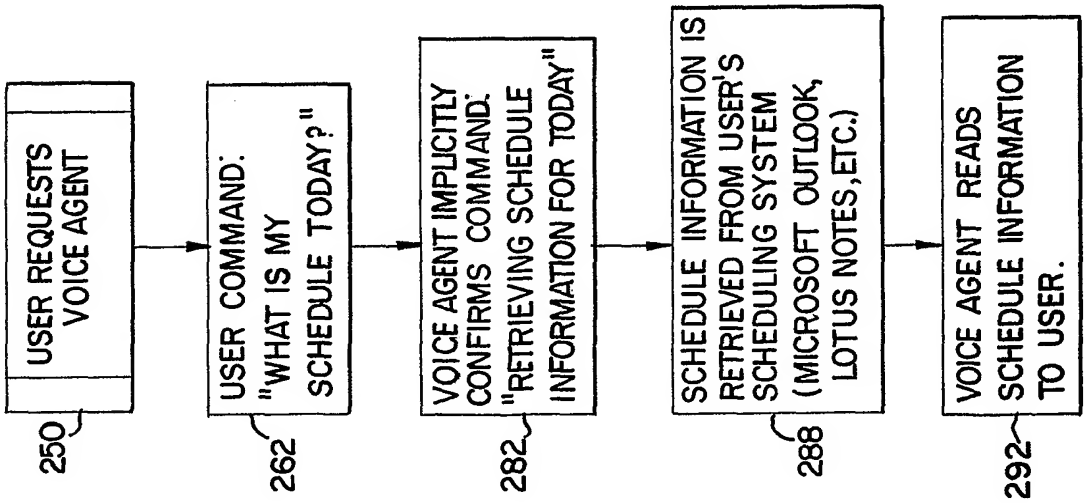


FIG. 13

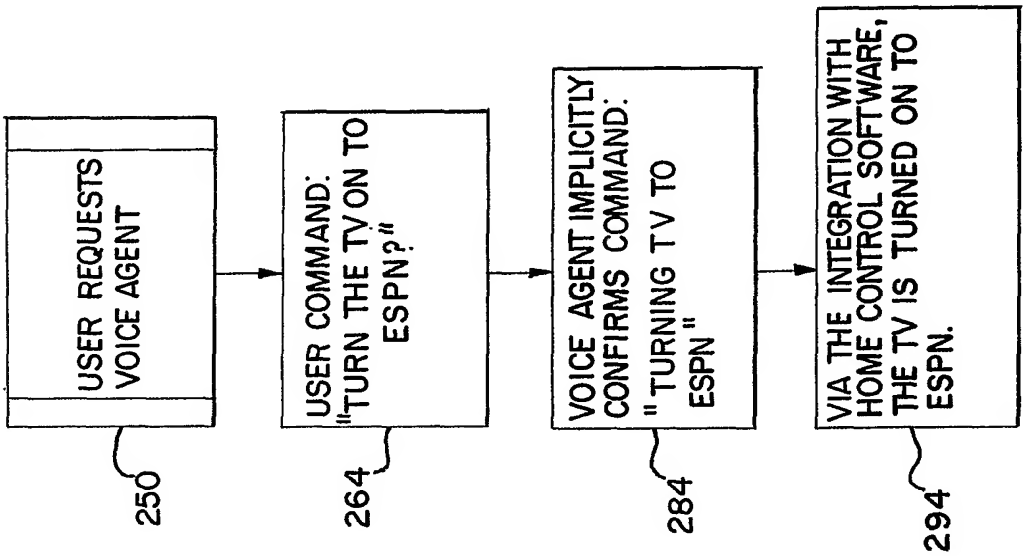
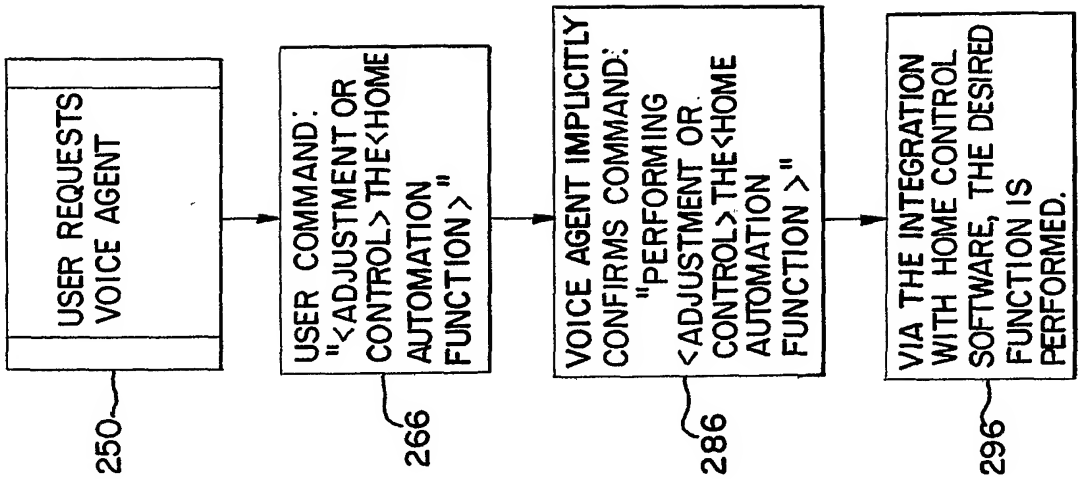


FIG. 14



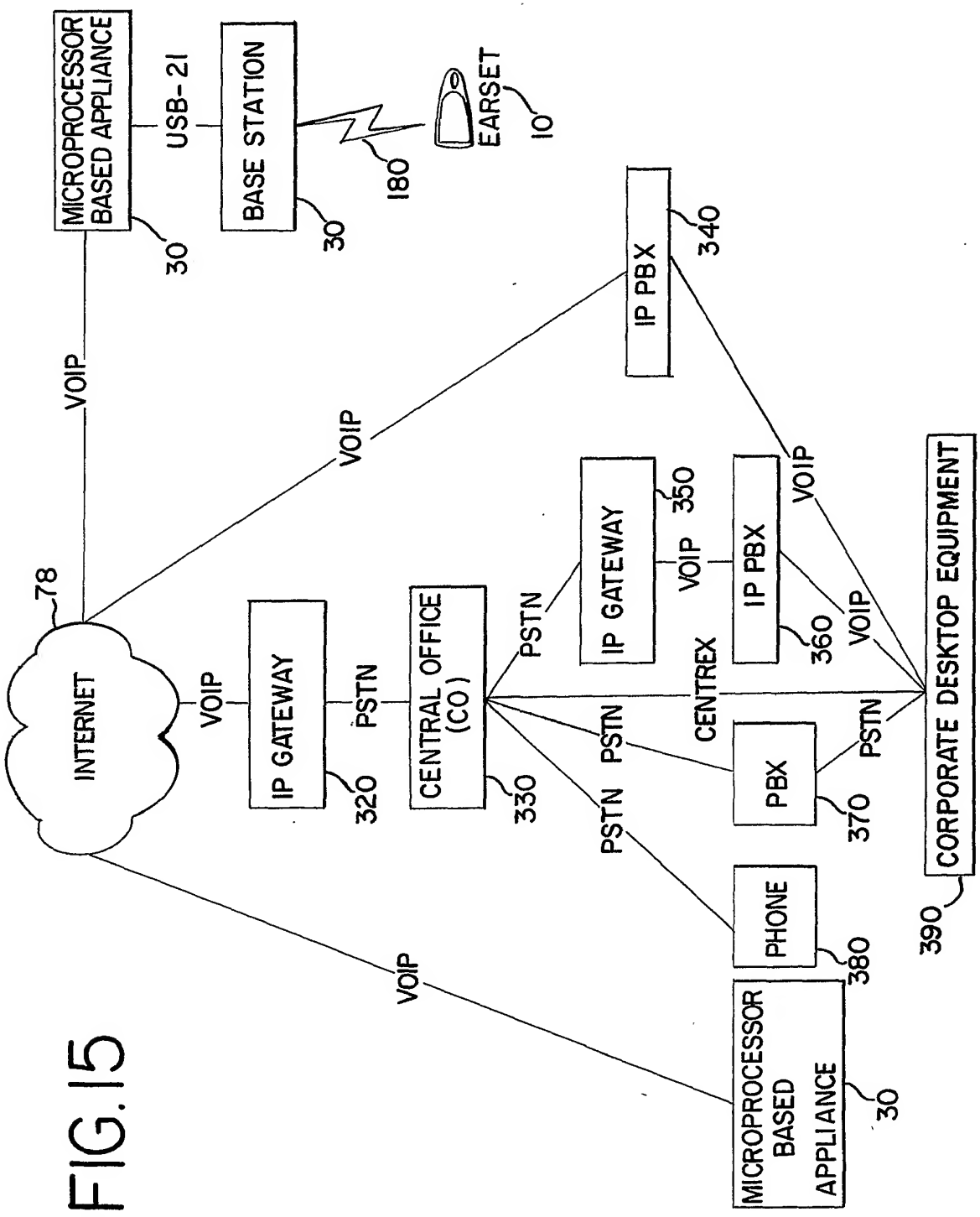

MICROPROCESSOR BASED APPLIANCE
USB-21
BASE STATION
EARSET
10
180
30
30
78
INTERNET
VOIP
VOIP
IP GATEWAY
320
PSTN
CENTRAL OFFICE (CO)
330
PSTN
IP GATEWAY
350
VOIP
IP PBX
340
VOIP
IP PBX
360
CENTREX
VOIP
PSTN PSTN
PBX
370
PSTN
PHONE
380
VOIP
MICROPROCESSOR BASED APPLIANCE
30
CORPORATE DESKTOP EQUIPMENT
390

FIG. 15